

**Е. В. БУРЦЕВА, А. В. ПЛАТЁНКИН,
И. П. РАК, А. В. ТЕРЕХОВ**

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ И СИСТЕМЫ



**Тамбов
Издательский центр ФГБОУ ВО «ТГТУ»
2024**

Министерство науки и высшего образования Российской Федерации

**Федеральное государственное бюджетное образовательное
учреждение высшего образования
«Тамбовский государственный технический университет»**

**Е. В. БУРЦЕВА, А. В. ПЛАТЁНКИН,
И. П. РАК, А. В. ТЕРЕХОВ**

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ И СИСТЕМЫ

Утверждено Ученым советом университета в качестве учебного пособия
для студентов 2, 3 курсов, обучающихся по направлению подготовки
09.03.03 «Прикладная информатика», всех форм обучения

Учебное электронное издание



Тамбов
Издательский центр ФГБОУ ВО «ТГТУ»
2024

УДК 004(075.8)
ББК з973я73
Б90

Рецензенты:

Кандидат технических наук, доцент,
заведующий кафедрой «КРЭМС» ФГБОУ ВО «ТГТУ»
Н. Г. Чернышов

Кандидат технических наук,
технический директор филиала ПАО «МТС» г. Тамбов
С. Б. Ушанёв

Бурцева, Е. В.

Б90 Информационные технологии и системы [Электронный ресурс] : учебное пособие / Е. В. Бурцева, А. В. Платёнкин, И. П. Рак, А. В. Терехов. – Тамбов : Издательский центр ФГБОУ ВО «ТГТУ», 2024. – 1 электрон. опт. диск (CD-ROM). – Системные требования : ПК не ниже класса Pentium II ; CD-ROM-дисковод ; 3,7 Мб ; RAM ; Windows 95/98/XP ; мышь. – Загл. с экрана.

ISBN 978-5-8265-2776-4

Представлены понятия и классификация информационных технологий, информационные технологии рабочего места пользователя, технологии баз данных, искусственные нейронные сети и методы машинного обучения. Рассмотрены справочно-правовые информационные системы и интернет-ресурсы сети, а также правовые основы защиты персональных данных.

Предназначено для студентов 2, 3 курсов, обучающихся по направлению подготовки 09.03.03 «Прикладная информатика», всех форм обучения, а также может быть полезно для тех, кто хочет получить базовые знания в области информационных технологий.

УДК 004(075.8)
ББК з973я73

*Все права на размножение и распространение в любой форме остаются за разработчиком.
Нелегальное копирование и использование данного продукта запрещено.*

ISBN 978-5-8265-2776-4

© Федеральное государственное бюджетное образовательное учреждение высшего образования «Тамбовский государственный технический университет» (ФГБОУ ВО «ТГТУ»), 2024

ВВЕДЕНИЕ

Роль информационных технологий и систем в развитии современного общества трудно переоценить, они задействованы во всех сферах человеческой деятельности. Их использование и развитие становятся все масштабнее и способствуют повышению эффективности технологического прогресса.

Перспективы развития информационных технологий (ИТ) требуют от пользователей соответствующих знаний и умений, повышая таким образом актуальность подготовки специалистов в ИТ-области.

Учебное пособие предназначено для подготовки студентов направления «Прикладная информатика в юриспруденции» и соответствует программе дисциплины «Информационные технологии и системы», входящей в обязательную часть дисциплин направления.

Материал, предложенный в электронном учебном пособии, представлен следующими темами:

- понятие информационной технологии;
- классификация информационных технологий;
- информационные технологии рабочего места пользователя;
- технология баз данных;
- классификация машинного обучения;
- проблемы при создании систем машинного обучения;
- пять направлений машинного обучения;
- искусственные нейронные сети;
- справочные правовые системы;
- некоторые вопросы информационной безопасности

и направлен на формирование знаний назначения и видов информационных технологий и информационных систем, способностей понимать принципы работы современных информационных технологий и программных средств, в том числе отечественного производства, и использовать их при решении задач профессиональной деятельности.

Пособие может быть полезно лицам, изучающим или осваивающим материал в области информационных технологий и систем.

1. ВВЕДЕНИЕ В ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

1.1. ПОНЯТИЕ ИНФОРМАЦИОННОЙ ТЕХНОЛОГИИ

Общеизвестно, что информационные технологии являются сегодня наиболее эффективным инструментом для работы с информацией. Определимся что такое *«информационная технология»*. Данное понятие включает в себя два слова: «информация» и «технология».

Словарь лингвистических терминов определяет понятие *«технология»* как какое-либо *искусство, мастерство, умение*, а именно – «совокупность методов, процессов и средств, используемых при производстве, изготовлении чего-либо» [32]. *Процесс* – представляет собой алгоритм действий, точное выполнение которых приводит к достижению поставленной цели, например, к написанию реферата (см. рис. 1). Чтобы реализовать процесс необходимы финансовые средства, определенные предметы труда и использование совокупности рациональных действий, т.е. методов исполнения данного процесса, пример которых представлен на рис. 2.

Чтобы понять, что такое *«информация»*, рассмотрим следующие определения. Одно из них дает закон об информации, информационных технологиях и о защите информации: *«информация – сведения (сообщения, данные) независимо от формы их представления»* [23, ст. 2]. Второе предлагают авторы учебника по информатике *«информация – это динамический объект, образующийся в ходе информационного процесса»* или *«продукт взаимодействия данных и методов, рассмотренный в контексте данного взаимодействия»* [11, с. 28].



Рис. 1. Процесс написания реферата

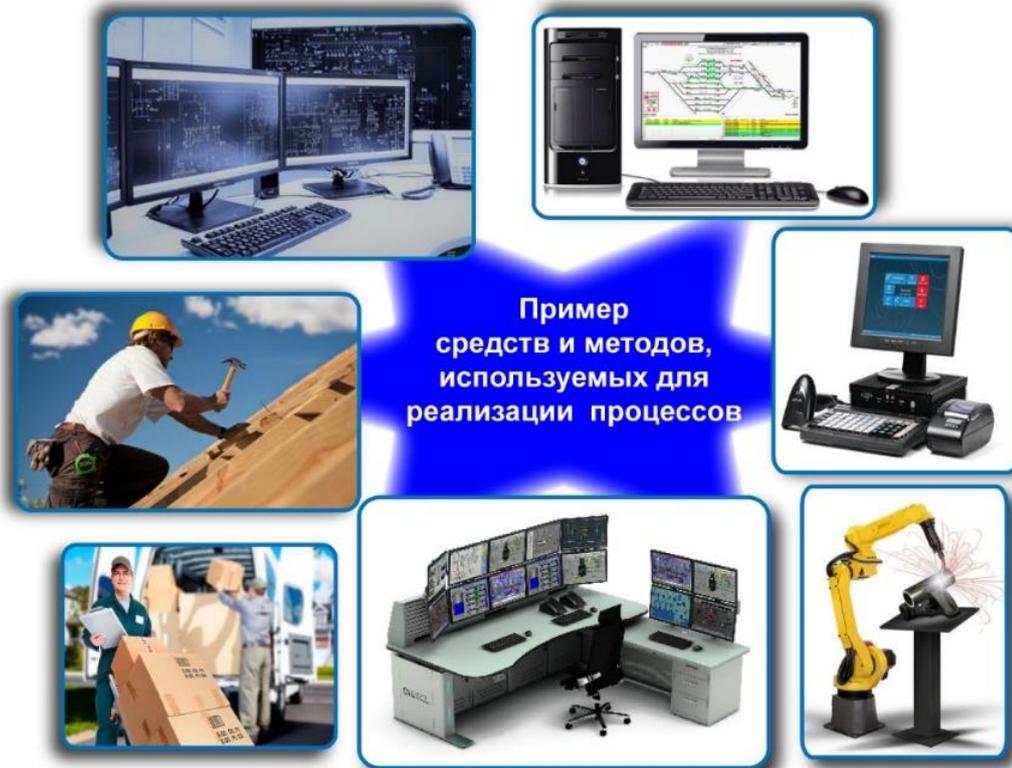


Рис. 2. Средства и методы, используемые для реализации процесса

Таким образом, *информационная технология* (ИТ) – это процесс, использующий совокупность средств и методов сбора, обработки и передачи первичной информации для получения информации нового качества [12, с. 15]. Данные авторы предлагают и более современное определение понятия «*информационная технология*» – «совокупность четко определенных целенаправленных действий персонала по переработке информации на компьютере» [19, с. 152]. Или, как определяет Федеральный закон «об информации, информационных технологиях и о защите информации», – это процессы, методы поиска, сбора, хранения, обработки, предоставления, распространения информации и способы осуществления таких процессов и методов [23, ст. 2].

Современные информационные технологии состоят из комплекса аппаратно-программных средств и системы организационно-методического обеспечения. В качестве технических (аппаратных) средств используются компьютерная техника и телекоммуникационное оборудование, а также внешнее дополнительное компьютерное оборудование. К программным средствам относятся системные и служебные программы и прикладное программное обеспечение.

В зависимости от инструментов, используемых для работы с информацией, информационные технологии подразделяются на следующие виды (см. рис. 3):

- «ручные» – в процессе работы с информацией не используются какие-либо приспособления, кроме ручки, карандаша, пера и т.п., обычной человеческой речи;



Рис. 3. Виды инструментария информационных технологий

- «механические» – для работы с информацией применяются механические печатные машинки, стационарные телефоны, средства доставки почты и т.д.;
- «электрические» – в данном случае информация обрабатывается с помощью электрических печатных машинок, копировальных аппаратов, диктофонов и т.д.;
- «электронные» – основным инструментарием являются большие электронно-вычислительные машины;
- «компьютерные» – в качестве инструмента используются персональные компьютеры, для которых создано большое количество программ, для решения разного класса задач.

Целью информационных технологий является производство информации для ее анализа человеком и принятия на основе результатов анализа решения по выполнению какого-либо действия. При этом основным преимуществом использования ИТ является снижение трудоемкости работы с информационными ресурсами и значительная экономия затрачиваемого на это времени. Последний показатель предлагается в качестве универсального количественного критерия эффективности информационных технологий.

В качестве основных правил работы информационных технологий следует отметить: диалоговый режим работы с компьютером, обеспечивающий свободную, т.е. не строго закреплённую последовательность выполнения операций обработки данных, возможность параллельного решения нескольких задач; гибкость процесса изменения данных и алгоритмов постановки задач;



Рис. 4. Требования к информационным технологиям

интегрированность с различными программными и аппаратными устройствами, все это обеспечивает уникальное свойство ИТ – умение свободного распространения знаний.

Современные информационные технологии должны отвечать требованиям, представленным на рис. 4.

Также в качестве особенностей информационных технологий, объясняющих их ролью в работе с информационными ресурсами, является стремительное развитие ИТ. Ежегодно появляются новые технологии, с использованием последних концепций и методов организации данных и работы с ними, в том числе на основе искусственного интеллекта, форм взаимодействия пользователей с аппаратно-программными средствами.

1.2. КЛАССИФИКАЦИЯ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ

Так как информационные технологии разрабатываются и используются для автоматизации разнообразных задач, в зависимости от назначения значительно различаются и структура информационных технологий, выполняемые функции, аппаратные средства, а также способы и методы обработки данных. В зависимости от функций, реализуемых информационной технологией, выполняемой задачи, а также ряда других признаков осуществляется классификация ИТ.

Видов классификации ИТ достаточно много. При этом многие информационные технологии могут одновременно относиться к разным классам.

Некоторые виды классификации ИТ представлены на рис. 5.



Рис. 5. Классификация информационных технологий

Обеспечивающие (базовые) информационные технологии – это инструментарий для решения всевозможных задач в различных предметных областях. Базовые технологии представлены комплексом технических и программных средств, наделенных широкими возможностями для работы с информацией, например, для ее извлечения, систематизации, обработки, передачи, интеграции и т.п. При этом базовые технологии выполняют не весь информационный процесс целиком, а только его часть, являются отдельной компонентой технологии реализации информационного процесса и служат для максимально эффективного выполнения этого отдельного фрагмента [14].

Функциональные (или предметные) ИТ – технологии, используемые при решении задач в различных сферах человеческой деятельности, – в медицине, строительстве, управлении, науке и т.п.

Однако деление ИТ на базовые и предметные, как и на многие другие, условно, так как одна и та же технология может быть и базовой, и предметной, когда базовую технологию приспособляют для решения задач в какой-либо предметной области.

Процессе объединения базовых информационных технологий на основе предметных с использованием соответствующих средств реализации этих технологий возможен только при приведении всех ИТ к единому стандартному интерфейсу [14].

Пакетные информационные технологии. Работа данных технологий обеспечивает непрерывным процессом обработки информации по введенному алгоритму без вмешательства пользователя. После ввода в компьютер исходных данных и программы для решения поставленной задачи, которые с учетом соответствующих правил объединяются в пакет, программа запускается и,

пока продолжается обработка пакета, прервать процесс нельзя. Результат виден только после выполнения последней операции. В качестве примера пакетной обработки данных можно привести процесс работы электронной почты, программы генерирования отчетности и другие.

Диалоговые информационные технологии. К ним относятся практически все современные технологии. Работа в диалоговом режиме предоставляет пользователю удобство и комфорт, а также неограниченные возможности использования информационных ресурсов системы в реальном масштабе времени. В процессе работы пользователю предоставляется широкая система справок и подсказок, вся информация, необходимая для решения задач и принятия решений. Диалоговый режим работы основан на прерываниях. Получив сигнал по линии прерывания, процессор приостанавливает текущую работу по программе, сохраняя при этом все временные данные, и переходит к новой задаче пользователя, которую также можно прервать, и так далее. Таким образом, диалоговый режим отличается свободой выполнения операций обработки данных, а не предполагает жестко закрепленной последовательности действий [14].

Сетевые информационные технологии – это базовые технологии интернет: Web-технологии, локальные, корпоративные, глобальные и комбинированные вычислительные сети, телекоммуникации, открытые системы, технологии поддержки распределенных вычислений и т.п. Они предназначены для обеспечения доступа к удаленным на любые расстояния, т.е. территориально распределенным ресурсам компьютерных сетей с помощью телекоммуникационных средств. С помощью сетевых технологий пользователи получили возможность работы с информацией, хранящейся на других компьютерах, а также выполнения удаленно одной задачи одновременно.

ИТ на базе локальной сети. Локальная сеть представляет собой систему взаимосвязанных компьютеров, объединенных либо на ограниченной территории, либо в рамках одной организации. Они обеспечивают надежное взаимодействие всех пользователей сети, предоставляя им возможность использования ресурсов сети (программно-аппаратными, информационными), совместной работы над одной задачей, свободной передачи, хранения и обработки информации.

Информационные технологии на базе многоуровневых сетей (сетей, представленных иерархической структурой) обеспечивают решение определенных задач на предназначенном для этого уровне, разграничивая таким образом доступ пользователей к ресурсам сети в зависимости от степени важности решаемых задач и реализуемых на каждом уровне функций управления.

ИТ на базе распределенных сетей. Задачей данных информационных технологий является обеспечение глобального взаимодействия пользователей сети, имеющих единую информационную инфраструктуру, с гарантией точной передачи информации между территориально удаленными узлами.

Функционально-ориентированные технологии используют принцип алгоритмической декомпозиции, который заключается в разделении задач предметной области на множество иерархически подчиненных функций. Для реализации конкретной функции используется своя информационная технология и устанавливается строгий порядок исполнения действий. Системы, созданные с использованием функционально-ориентированного принципа, отличаются однонаправленностью, поэтому в случае изменения требований к системе нет возможности вернуться назад, можно только полностью перепроектировать систему, это является главным недостатком функционально ориентированного подхода [14].

Объектно-ориентированные информационные технологии используют принцип проектирования открытой системы в виде взаимосвязанных, но функционально независимых объектов, которые совместно, с требуемой эффективностью выполняют заданные функции системы. Информационные технологии, созданные на базе объектно-ориентированного подхода, обладают возможностью быстрой адаптации к вновь появляющимся задачам за счет набора специфических свойств (наследование и др.). Таким образом, значительно снижаются затраты на их разработку, внедрение и модификацию [14].

Данный вид технологий является наиболее эффективным и используется сегодня в большинстве информационных систем, особенно в системах управления производством.

Информационные технологии обработки данных предназначены для решения функциональных задач, по которым имеются все входные данные, известны алгоритмы и стандартные процедуры их обработки, т.е. для выполнения относительно несложных, стандартных операций с данными [10, 30].

Информационные технологии управления – предназначены для автоматизации труда тех работников предприятий, в обязанности которых входят функции принятия каких-либо управленческих решений.

Информационные технологии автоматизации офисной деятельности – реализуют организацию и поддержку работы офиса, автоматизируют процессы работы и документами от их создания, согласования до пересылки, обеспечивают контроль управления.

Информационные технологии поддержки принятия решений – это технологии, основанные на использовании математического аппарата, искусственного интеллекта и предназначенные для обработки и анализа больших объемов данных, результатом которых являются отчеты, прогнозы, выводы, необходимые для принятия обоснованных решений [2].

Экспертные системы – это особый вид автоматизированных систем, основанных на технологиях искусственного интеллекта, и состоящих из базы знаний, набора сформулированных экспертами эвристических правил и механизма вывода. Подобные системы, на основе встроенных правил, элементов

анализа и интеллекта, а также опыта, накопленного в конкретной области знаний, способны распознать ситуацию, поставить диагноз, сформулировать решение предоставленных пользователем фактов или дать рекомендацию для выбора действия, аналогично тому, как это делают эксперты при выработке умозаключений и логических выводов. Экспертные системы предназначены для решения задач, традиционно относящихся к области деятельности человеческого интеллекта, и являются одним из наиболее развивающихся классов интеллектуальных систем [2, с. 22]. Они составляют основу автоматизации труда специалистов-аналитиков.

Правильная классификация позволяет понять, оценить, грамотно разрабатывать и использовать информационные технологии в различных сферах деятельности человека.

2. ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ РАБОЧЕГО МЕСТА ПОЛЬЗОВАТЕЛЯ

Как показывает статистика, с каждым годом в России растет процент населения, пользующегося какими-либо информационными технологиями. Например, к началу 2022 года процент использования Интернета составил 89% от всего населения страны [26]. Принятая национальная программа «Цифровая экономика Российской Федерации» повысила уровень использования информационно-коммуникационных технологий в деловой сфере, привела к повсеместному их проникновению «в отрасли хозяйства, общественную сферу и государственное управление» [20].

Все это вызывает потребность в новых ИТ-решениях, повышении эффективности, безопасности и удобства использования информационных технологий [35].

На рабочих местах пользователей могут использоваться как универсальные программные средства, так и специальные. К универсальным программным средствам относятся все программы, входящие в офисные пакеты: текстовые и табличные редакторы, редакторы презентаций, некоторые графические редакторы, а также все сервисы сети интернет, оптические системы распознавания символов, машинные переводчики, т.е. те программы, которые используются пользователями на любом рабочем месте или в быту. Специализированные информационные технологии – это программные и технические средства, создаваемые для решения задач в конкретной узкой области деятельности и предназначенные для решения определенной задачи, например, справочные правовые системы, их задача – обеспечение специалистов в области права и других работников (экономистов, студентов, преподавателей, сотрудников отделов кадров) достоверной нормативной правовой информацией.

2.1. АВТОМАТИЗИРОВАННОЕ РАБОЧЕЕ МЕСТО

Автоматизированное рабочее место (АРМ) – это «программно-технический комплекс автоматизированной системы, предназначенный для автоматизации деятельности определенной категории пользователей или определенного вида деятельности» [6]. Для создания АРМ сегодня, как правило, используются персональные компьютеры, включенные в вычислительную сеть организации, и другие необходимые электронные средства, а также требуемые для данного специалиста программные средства. В качестве программных средств АРМ может использоваться набор программ либо специализированная программа – автоматизированная информационная система или автоматизированное рабочее место, реализующая весь процесс выполнения большинства ежедневных задач определенного специалиста, что является основной целью любого АРМ [2].



Рис. 6. Принципы построения организационных форм обработки данных

Использование автоматизированных рабочих мест обеспечивает выполнение основных функций по накоплению, хранению, переработке и выдаче информации информационными технологиями, на пользователя возлагаются задачи по вводу исходной информации, а также задачи, требующие творческого подхода при принятии управленческих решений.

При создании автоматизированных рабочих мест следует руководствоваться следующими основными принципами, представленными на рис. 6 [37].

Принцип устойчивости означает, что автоматизированное рабочее место должно обеспечивать стабильную и устойчивую работу, защиту информации от потери. Для этого следует обеспечивать АРМ средствами автосохранения и восстановления данных в случае какого-либо сбоя, например, при отключении электричества, и/или возможностью дублирования информации.

Принцип системности заключается во взаимосвязи всех элементов рабочего места.

Гибкость АРМ означает его адаптированность к изменениям, регулярному развитию компьютерной техники и информационных технологий, т.е. возможность модернизации АРМ.

Принцип эффективности направлен на обеспечение оптимальной результативности функций, реализуемых автоматизированным рабочим местом. Кроме того, данный принцип заключается в том, что АРМ будет иметь понятный интерфейс, обеспечит быструю обработку запросов, простоту обслуживания, а работа с ним будет легкой и комфортной для пользователя.

Автоматизированные рабочие места можно классифицировать по нескольким признакам, представленным на рис. 7.



Рис. 7. Виды автоматизированных рабочих мест

В состав автоматизированных рабочих мест входят различные виды обеспечения, представленные на рис. 8.



Рис. 8. Виды обеспечения автоматизированных рабочих мест

Техническое обеспечение АРМ представлено комплексом средств для оснащения рабочего места специалиста. Его основу составляют персональные компьютеры, состав и мощность которых должны обеспечить решение задач конкретного специалиста, а также необходимый перечень периферийных устройств.

Информационное обеспечение представлено информационной базой АРМ, определяет ее структуру и правила информационных связей.

Математическое обеспечение – это совокупность математических методов, моделей и алгоритмов обработки информации, необходимых для реализации целей и задач АРМ.

Программное обеспечение направлено на реализацию целей и задач АРМ, а также нормальное функционирование комплекса технических средств. Оно должно осуществлять функции, представленные на рис. 9.

Лингвистическое обеспечение представлено языковыми средствами АРМ, которые включают язык общения пользователя с компьютером, языки запросов и отчетов, языки определения и управления данными. Лингвистическое обеспечение во многом определяется спецификой предметной области и зависит от требований, которые предъявляются к полноте и точности передачи информации, а также от требований к унифицированности языка, простоты его освоения и использования [14, с. 34].

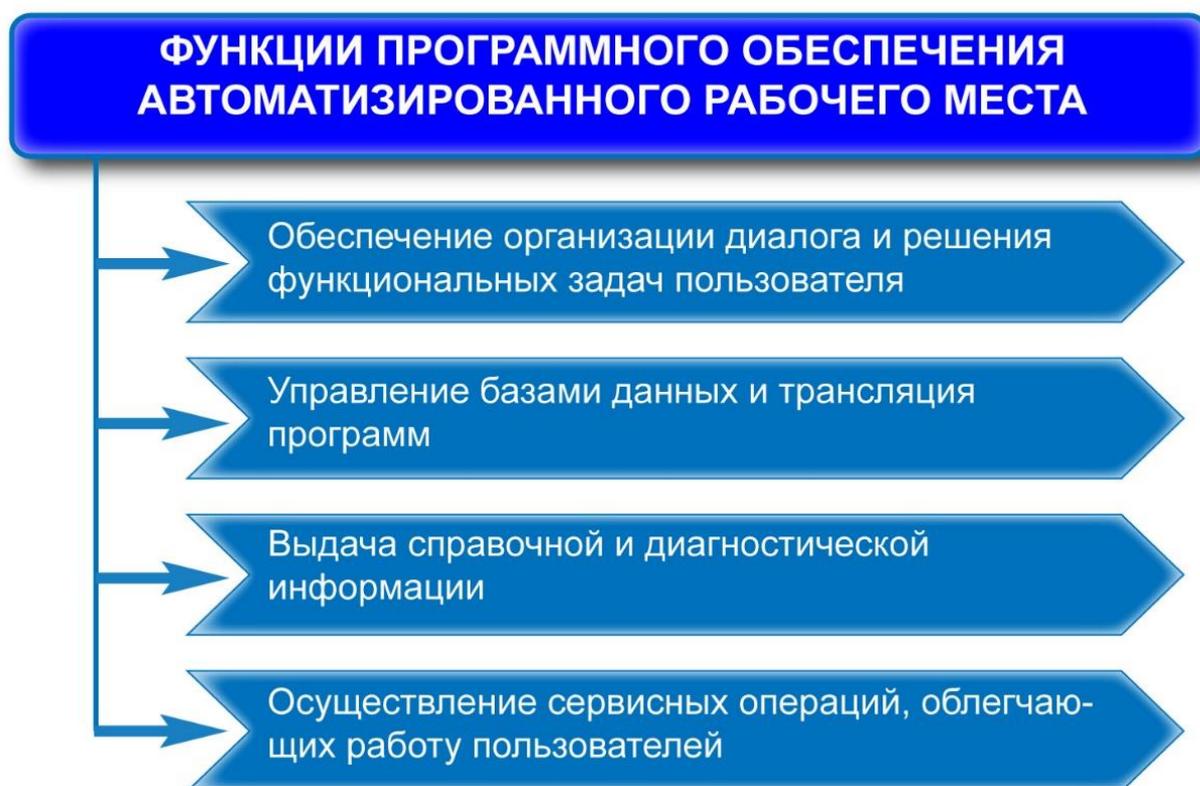


Рис. 9. Функции программного обеспечения АРМ

Технологическое обеспечение АРМ – это совокупность используемых методов, которые устанавливают последовательность этапов выполнения операций, процедур при решении установленных для конкретной АРМ задач, например, процедуру поиска необходимых данных.

Организационное обеспечение включает комплекс документов, регламентирующих деятельность специалистов в процессе разработки и эксплуатации АРМ. В документах, в частности, регламентируются:

- функции и задачи каждого специалиста;
- условия взаимодействия сотрудников;
- правила выполнения технологических операций автоматизированной обработки информации.

Методическое обеспечение АРМ – это вся справочная информация, включающая сведения об АРМ – руководство пользователя, перечень функций, выполняемых АРМ, рекомендации по внедрению и эксплуатации и др.

Эргономическое обеспечение содержит информацию об условиях, создающих максимально комфортную среду для использования АРМ. Сюда входит оснащение рабочего места необходимым оборудованием, в том числе мебелью, обеспечивающих удобную и комфортную работу пользователя.

Правовое обеспечение АРМ – это совокупность правовых норм, определяющих вопросы создания, юридический статус и условия использования АРМ, регламентирующих порядок работы с информацией и ее защиты [14, с. 34].

2.2. ЭЛЕКТРОННЫЙ ОФИС

Электронный офис представляет собой офисную работу, реализуемую с использованием информационных технологий, которые обеспечивают автоматизацию документооборота и управления организацией.

На рисунке 10 представлен перечень стандартных типовых процедур обработки документов с помощью компьютерных технологий.

Электронный офис может быть представлен комплексом необходимых программных и технических средств, автоматизирующих работу специалистов офиса.

Аппаратные средства электронного офиса – это электронные устройства, как правило, компьютерные, необходимые для решения офисных задач. Они подразделяются на основные и дополнительные.

В настоящее время основными аппаратными средствами для выполнения офисной работы являются персональные компьютеры, как правило, работающие в сети, и набор необходимых периферийных устройств для ввода и вывода информации (см. рис. 11).

Остальные технические средства, используемые в офисе и не нуждающиеся в подключении к компьютеру, называются дополнительными. Они обеспечивают сотрудникам комфортные условия работы, например, ксероксы, шредеры, телефоны и факсы и др.



Рис. 10. Процедуры обработки документов в электронном офисе



Рис. 11. Аппаратные средства электронного офиса



Рис. 12. Программные средства электронного офиса

Программные средства электронного офиса – это пакеты офисных прикладных программ, которые позволяют автоматизировать основные процедуры обработки информации, а также специализированные информационные системы, предназначенные для автоматизации конкретных офисных задач.

Программные средства, используемые для автоматизации офисной деятельности, можно представить в следующем виде (см. рис. 12).

В качестве программных средств, используемых для работы с документами, следует отметить все прикладные программы офисных пакетов, это текстовые, табличные и графические процессоры (или редакторы), редакторы презентаций, публикаций и т.д. (см. рис. 13).

Кроме того, существуют и самостоятельные программные продукты, используемые в офисной работе, в частности, оптические системы распознавания символов, машинные переводчики и другие.

Организаторы работ – это информационные системы управления процессами и ресурсами организации: операционными задачами и поручениями, проектами, договорами, финансами, персоналом и активами.



Рис. 13. Характеристика основных пакетов программ формирования и редактирования документов

Такие программы автоматизируют большинство процессов организации [31, 33]:

- документооборот;
- полный цикл договорной работы;
- планирование и постановку задач и поручений подчиненным;
- формирование аналитических отчетов;
- управление проектами на всех стадиях;
- планирование совещаний и заседаний;
- контроль исполнения принятых решений;
- управление финансово-хозяйственной деятельностью,
- взаимоотношения с клиентами за счет погружения всей деятельности единую информационную рабочую среду (с возможностью голосового управления).

Системы электронного документооборота – это программы с полным набором инструментов для управления документооборотом и делопроизводством в организации [28], автоматизирующие следующие функции при работе с документами:

- подготовка, согласование, обсуждение проектов документов и их утверждение;
- редактирование документов, в том числе совместное;
- получение и отправка документов по почте;
- маршрутизация документов;
- сканирование бумажных версий документов;
- автоматическая регистрация электронных версий документов в регистрационной карточке документа;
- контроль исполнения поручений;
- передача документов в архив;
- ведение электронных архивов;
- составление реестров документов;
- подготовка отчетов;
- и многое другое.

В последнее время многие программные продукты для выполнения офисной работы стали создаваться для мобильных устройств и в качестве онлайн приложений.

3. ТЕХНОЛОГИИ БАЗ ДАННЫХ

3.1. РАЗВИТИЕ БАЗ ДАННЫХ

Базы данных (БД) являются одним из важнейших компонентом большинства современных информационных систем (ИС) и играют ключевую роль в организации, хранении и управлении больших объемов данных.

Базы данных используются повсеместно и во всех сферах деятельности человека. Финансовые, медицинские, учебные, научные и прочие учреждения и организации так или иначе используют в своей деятельности системы хранения и обработки данных [1].

Предпосылкой для появления баз данных явилось создание в конце 50-х годов прошлого века централизованной системы управления файлами, которая отвечает за создание, уничтожение, организацию, чтение, запись, модификацию и перемещение файловой информации, а также за управление доступом к файлам и за управлением ресурсами, которые используются файлами. В настоящее время система управления файлами является основной подсистемой любой современной операционной системы (ОС).

В последующем стало очевидным, что совместно используемые данные должны обладать специфическими свойствами, в частности: независимость данных, отсутствие дублирования и противоречивости, контроль прав доступа к данным, эффективная техника доступа к данным и многое другое.

Другим фактором, повлиявшим на развитие баз данных, стало появление и относительно широкое распространение в начале 1960-х гг. магнитных дисков. Эти запоминающие устройства имели достаточно большой для того времени объем памяти и возможность доступа к произвольным участкам памяти, что позволило создавать сложные структуры долговременно хранимых данных [22].

В начале вопрос хранения данных разработчики ИС решали самостоятельно с помощью функций ОС или напрямую обращаясь к устройствам ввода-вывода. Однако ОС способна предоставить только функции по работе с файлами, а вопросы организации хранения записей внутри файла, поиска данных, проверки ограничений для записи ее средствами решить невозможно. Кроме того, при одновременном доступе нескольких пользователей к одним и тем же данным необходимы дополнительные механизмы, позволяющие централизованно управлять этим процессом. Эти и ряд других причин привели к созданию отдельного класса программного обеспечения – *систем управления базами данных* (СУБД), а сами хранилища информации, которые работали под их управлением, назывались *базами* или *банками данных*.

В развитии БД можно выделить несколько этапов (периодов).

Первый этап (период становления) принято считать с начала 60-х по первую половину 70-х годов XX века. В это время появляется термин «база данных» и разрабатываются первые СУБД.

Появление больших ЭВМ (типа IBM 360/370, ЕС ЭВМ) с магнитными дисками привело к появлению в середине 60-х гг. первых СУБД, в качестве примера можно привести системы IMS (англ. Information Management System, Система управления информацией) фирмы IBM и IDS (англ. Integrated Data Store, Интегрированное хранилище данных) компании General Electric. СУБД IMS поддерживала иерархическую (англ. hierarchical) структуру данных, а IDS – сетевую (англ. network).

Первые БД хранились во внешней памяти центральной ЭВМ и обработка данных выполнялась в пакетном режиме. Интерактивный режим доступа, осуществляемый с помощью терминалов, не имеющих собственные вычислительные ресурсы, обеспечивал только ввод данных для центральной ЭВМ и вывод результатов [21].

Программы доступа к БД писались на обычных языках программирования. СУБД этого периода работали с централизованной базой данных в режиме распределенного доступа, при этом функции управления распределением ресурсов выполнялись ОС.

Следующий этап БД (период развития) связан с появлением реляционной модели данных. В 1970 году Э. Ф. Коддом, работавший в исследовательской лаборатории фирмы IBM, опубликовал статью «A Relational Model of Data for Large Shared Data Banks» (Реляционная модель данных для больших общих банков данных), в которой впервые прозвучал термин «реляционная модель». Проводятся серьезные исследования теоретических и прикладных вопросов реляционной модели, которые привели к созданию формальной теории баз данных.

Большой вклад на развитие теории и практики реляционных систем во всем мире оказала разработанная в конце 1970-х гг. в исследовательской лаборатории фирмы IBM, СУБД System R. Хотя System R не получила коммерческого успеха, однако она продемонстрировала жизнеспособность реляционной модели данных, а также в ней был реализован SQL (англ. Structured Query Language, Структурированный язык запросов), который стал стандартом на язык запросов для реляционных БД.

В 1980-е годы (период зрелости) реляционная модель данных уже обрела полное теоретическое обоснование. В это время появляется большое количество различных реляционных СУБД (например, Oracle, Informix), которые находят широкое применение в различных сферах человеческой деятельности. Реляционные системы практически вытеснили с мирового рынка ранние СУБД иерархического и сетевого типа [15].

Начиная с середины 1980-х годов (постреляционный период) проводятся интенсивные исследования, связанные с моделями данных, отличных от классической реляционной. Появляются СУБД, основанные на объектно-ориентированной парадигме, где данные моделируются в виде объектов, их атрибутов, методов и классов, что позволяет работать с объектами БД так же, как с объектами в объектно-ориентированных языках программирования.

В начале 2000-х годов появились новые подходы к хранению данных, среди которых можно выделить документо-ориентированную модель. Катализатором развития документо-ориентированных БД стало появление распределенных систем и технологий обработки больших данных (англ. Big Data).

В это же время появляется понятие NoSQL (англ. Not only SQL, не только SQL). NoSQL-базы данных предназначены для более крупных объемов данных, которые сложно структурировать с использованием реляционного подхода. Они позволяют быстрее обрабатывать более масштабные наборы данных, имеющих различную структуру.

Также следует отметить, что на развитие технологий БД сильное влияние оказало появление персонального компьютера (ПК), локальных сетей (ЛС) и Интернета.

Появление в конце 70-х годов прошлого века ПК привело к тому, что компьютерная техника получила широкое распространение и стала использоваться для решения различных задач. В это время появляются, так называемые, настольные (англ. desktop) СУБД. Такие СУБД, как правило, могли обеспечить к БД только монопольный доступ, т.е. в определенный момент с БД могло работать только один пользователь.

Наличие большого количества настольных СУБД на рынке потребовало от их разработчиков постоянно расширять их возможности и улучшать пользовательский интерфейс, а также разработки форматов хранения данных и методов экспорта-импорта данных из одного формата в другой.

В качестве примера СУБД такого класса можно привести Dbase, Paradox, FoxPro, Clipper.

Появление ПК также привело к тому, что многие компании вместо одного большого компьютера (мейнфрейма) стали приобретать несколько ПК. Со временем появилась потребность объединять компьютеры между собой, так в начале 1980-х годов стали появляться ЛС.

Сначала в ЛС использовались сетевые СУБД, основанные на файл-серверной архитектуре. В этом случае БД хранилась на специально выделенном для этого компьютере (файл-сервере), а на клиентском компьютере устанавливалась копия СУБД. Файл-серверные СУБД работают в однопользовательском режиме и могут вызывать существенную перегрузку сети, так как вся обработка данных происходит на стороне клиента, что требует от СУБД пересылать весь файл БД целиком. Также следует отметить невысокую надежность систем на основе файл-серверной архитектуры, так как сбой на каком-нибудь этапе может привести к потере данных.

Следующим шагом в развитие БД стало появление клиент-серверной архитектуры, которая обеспечила многопользовательский режим доступа к данным и высокую надежность систем. Такая модель предполагает распределение нагрузки и заданий между теми, кто предоставляет услуги («сервером») и теми, кто их использует («клиент»). На сервере располагается БД, он обрабатывает запросы от клиентов и отправляет необходимые им данные. Клиент

формирует и отправляет запрос к серверу и выполняет окончательную обработку полученных от сервера данных.

Стремительное развитие ИТ (вычислительной техники, локальных сетей, Интернета) привело к тому, что большой объем данных располагается на различных компьютерах, территориально удаленных друг от друга, функционируют под управлением разнообразных ОС, а доступ к ним осуществляется разнородным программным обеспечением. Встают задачи согласованности данных, хранящихся и обрабатываемых в разных местах, но логически друг с другом связанных, и параллельной обработкой данных. Это приводит к появлению распределенных СУБД.

Необходимость поддержки многопользовательской работы с БД и возможность децентрализованного хранения данных потребовали развития в СУБД средств администрирования БД.

3.2. КЛАССИФИКАЦИЯ БАЗ ДАННЫХ

Наиболее распространенный способ классификации БД основан на используемой в них модели данных. В настоящее время выделяют два основных класса БД: реляционные (SQL) и нереляционные (NoSQL). В свою очередь, существуют различные виды NoSQL БД: типа «ключ-значение», документные, графовые, колоночные, объектно-ориентированные, поисковые и т.д. Так же следует отметить, что первые БД имели иерархическую и сетевую модели данных.

Реляционные БД используются, когда необходимо обрабатывать структурированные данные, а нереляционные эффективны при работе с неструктурированными или слабоструктурированными данными.

Иерархические БД. Иерархическая модель данных использует представление данных в виде древовидной (иерархической) структуры, состоящей из объектов различных уровней. Каждый объект представляется в виде определенной сущности, которая может иметь дочерние и родительский элементы. Родительский элемент может иметь несколько дочерних элементов, а дочерний только одного предка (см. рис. 14).

Сетевые БД. Сетевая модель данных является модификацией иерархической. В ее основе лежит структура в виде графа, которая, в отличие от иерархической модели, позволяет потомку иметь более одного предка [1].

Реляционные БД. Реляционную модель принято считать классической и на данный момент она является самой распространенной.

В основе реляционной модели лежит математическая теория множеств, что подразумевает представление информации в виде двумерных таблиц, имеющих предопределенные отношения друг с другом (см. рис. 15).

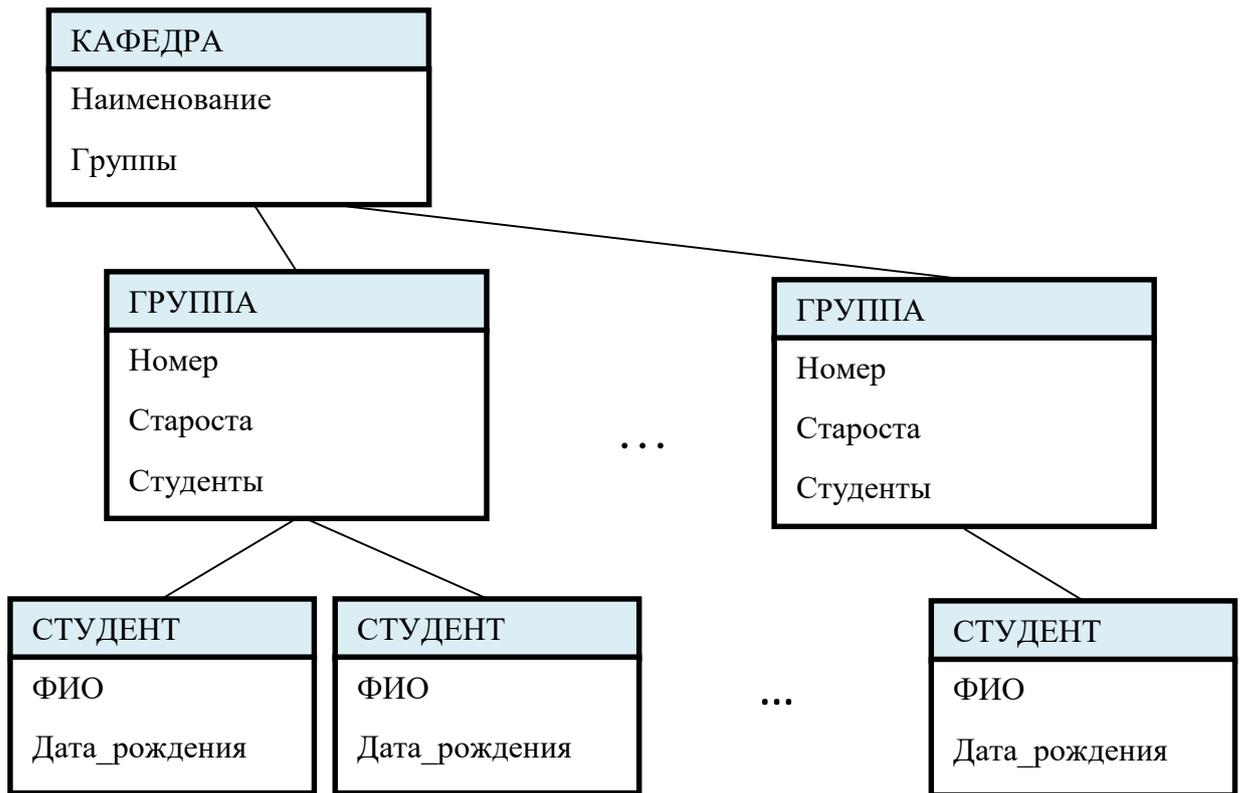


Рис. 14. Иерархическая БД



Рис. 15. Реляционная БД

Структура реляционной БД заранее известна, т.е. определены таблицы (они имеют фиксированное количество столбцов, каждый столбец имеет имя и тип данных) и связи между ними, которые осуществляются через общий для двух таблиц столбец. В каждой таблице должно быть хотя бы одно ключевое поле, содержимое которого уникально для любой записи в этой таблице.

Реляционные СУБД управляются с помощью языка SQL.

К реляционным СУБД относятся, например, Oracle Database, Microsoft SQL Server, PostgreSQL, Postgres Pro, MySQL, SQLite, JatoBa.

Объектно-ориентированные БД. В объектно-ориентированных БД, в отличие от реляционных, данные моделируются не в виде таблиц, а как объектов с атрибутами и методами.

Объектно-ориентированные БД предлагают разработчикам работать непосредственно с объектами и классами. Это упрощает процесс моделирования данных и позволяет создавать более гибкие и выразительные схемы для представления информации. Важными особенностями таких БД является возможность определения пользовательских типов данных и наличие механизма наследования, который позволяет создавать иерархические структуры данных.

В качестве примера объектно-ориентированных СУБД можно привести: db4o, DataFlex, ZODB, ObjectStore.

Объектно-реляционные БД. Объектно-реляционные БД объединяют в себе возможности реляционной и объектной моделей. Их возникновение объясняется тем, что реляционные БД хорошо работают со встроенными типами данных и в случае появления нового важного типа данных приходится либо включать его поддержку в СУБД, либо заставлять программиста самостоятельно управлять данными в приложении.

БД «ключ-значение». «Ключ-значение» (англ. key-value) представляет из себя простейший вариант хранения данных. В таких БД в таблице хранится уникальный ключ и связанное с ним значение, в качестве которого может быть любая информация. БД данного типа применяются, когда имеются большие наборы данных, по которым нужно быстро производить поиск, например, для хранения изображений, в качестве кэшей для объектов, в Big Data системах.

Самые известные СУБД типа «ключ-значение»: Oracle NoSQL Databases, Redis, MemcacheDB, Amazon DynamoDB, Berkeley DB, Riak.

Документно-ориентированные БД. Документно-ориентированные БД предназначены для хранения иерархической коллекции объектов (документов), причем в коллекцию могут входить документы разных форматов [9]. Хранилище имеет структуру дерева или леса. Деревья начинаются с корневого узла и может содержать несколько внутренних и листовых узлов. Листовые узлы содержат данные, которые при добавлении документа заносятся в индексы, это дает возможность даже при достаточно сложной структуре находить путь к искомым данным.

Основным инструментом по доступу и управлению данными выступает декларативный язык NoSQL.

Самые известные документные СУБД: CouchDB, MongoDB, Amazon DocumentDB, Berkeley DB XML.

Графовые БД. В графовых БД реализуется сетевая модель, где для отображения и хранения данных используются узлы и ребра графа. Узлы содержат информацию о некоей сущности, т.е. ее свойства, а ребра – об отношениях между узлами. Данный вид БД был разработан специально для того, чтобы хранить и обрабатывать сведения о взаимоотношениях между сущностями. Ребро может описывать отношения типа родитель-потомок, действия, отношения владения ресурсами и т.п. При этом количество и тип отношений, которые может иметь узел, не ограничено.

Графовые СУБД обеспечивают высокое быстродействие при решении задач, связанных с обходом графа или поиском подграфов, так как ребра также являются хранимыми данными. Такие СУБД применяются, например, в социальных сетях, дорожных картах, маршрутах общественного транспорта, в финансовой сфере для выявления мошенничества [9, 21].

Наиболее распространенные графовые СУБД: Open Source, Neo4j, Amazon Neptune, InfiniteGraph, InfoGrid, OrientDB, Titan.

Колоночные БД. В колоночных (столбцовых, столбцово-ориентированных) БД данные хранятся в разреженной матрице, строки и столбцы которой используются как ключи. Строки таблицы сортируются по ключу строк (первичному ключу таблицы). Столбцы объединяются в семейства, которые задаются как часть определения схемы таблицы, однако новые члены семейств могут добавляться по мере надобности. В таких БД, чтобы получить значение атрибута одного из объектов в таблице, не надо читать всю строку до нужного столбца, как в реляционной БД, необходимый атрибут будет прочитан сразу. Это позволяет ускорить чтение при больших объемах, поэтому они используются при работе с Big Data.

Колоночные БД содержат отметки времени, что позволяет использовать их для организации счетчиков, регистрации и обработки событий, связанных со временем, например, в системах биржевой аналитики, Интернете вещей, управления содержимым.

Примеры колоночных СУБД: SAP IQ, Vertica, ClickHouse, Google BigTable, Apache HBase, InfoBright, Apache Cassandra, ScyllaDB, Hypertable.

Поисковые БД. В поисковых БД осуществляется собственный оптимизированный подход к индексированию данных, позволяющий выполнять автоматизированный поиск по содержимому документа.

Такие системы подходят для быстрого полнотекстового поиска в различных источниках как структурированных, так и неструктурированных данных.

Самые известные поисковые СУБД: Apache Solr, Elasticsearch, Splunk.

Следует отметить, что это не полный перечень существующих разновидностей БД, технологии развиваются и периодически появляются новые виды БД для решения узкоспециализированных задач.

В зависимости от расположения БД их можно разделить на *локальные* и *сетевые (серверные)*.

Локальные БД предполагают систему хранения данных на одном компьютере, как правило, на сервере внутри организации.

Сетевые БД могут быть реализованы с использованием клиент-серверных или облачных технологий. Использование технологии «клиент-сервер» означает, что СУБД и БД размещены на одном сервере, к которому пользователи обращаются со своими запросами. Сетевые БД могут быть централизованными (БД хранится на одном компьютере) и распределенными (БД частично находится на разных серверах). Доступ к сетевым БД может осуществляться в том числе с использованием облачных технологий.

Также бывают встраиваемые СУБД – локальные СУБД, которые представляют собой отдельный модуль для управления данными внутри приложений. Обычно они написаны в виде библиотек для разных языков программирования.

По типу хранимых данных БД делят на документальные, фактографические и лексикографические. В свою очередь документальные подразделяются на библиографические, реферативные и полнотекстовые.

В **документальных** БД единицей хранения является какой-либо документ (например, закон) или ссылка, по которой можно обратиться к исходному документу. БД, в которой предусмотрено хранение полного текста документа, называется полнотекстовой, а библиографические и реферативные содержат ссылки.

В **фактографических** БД хранятся данные об объектах предметной области в виде «фактов» (например, биографические данные сотрудников, данные о выпуске товара и т.п.).

К **лексикографическим** БД относятся различные словари (классификаторы, многоязычные словари, словари основных слов и т. п.).

По характеру организации данных БД могут быть разделены на: неструктурированные, частично структурированные и структурированные.

Этот классификационный признак относится к данным, представленным в символьном виде. К **неструктурированным** БД могут быть отнесены базы, организованные в виде семантических сетей. **Частично структурированными** можно считать БД в виде обычного текста или гипертекстовые системы. **Структурированные** БД требуют предварительного проектирования и описания структуры БД. Только после этого БД такого типа могут быть заполнены данными. К таким относятся реляционные БД [22].

БД можно классифицировать по функциональному назначению, то есть по тому, для каких целей они предназначены и какие задачи они решают. По этому признаку можно выделить операционные, аналитические, распределенные БД, хранилища данных и т.д.

Операционные БД используются для хранения и обработки данных, связанных с операционной деятельностью организации. Они предназначены для поддержки операций, которые в организации выполняются постоянно, например, обработка заказов, учет товаров, управление клиентскими данными и т.д. Такие БД должны обеспечивать быстрый доступ к данным и обновление информации в режиме реального времени.

Аналитические БД используются для анализа больших объемов данных и выявления закономерностей, трендов и шаблонов. Они предназначены для поддержки принятия решений на основе данных и проведения различных аналитических операций, таких как агрегирование, сортировка, фильтрация и т.д. Такие БД обычно содержат исторические данные и предоставляют возможность проводить сложные аналитические запросы.

Распределенные БД состоят из нескольких физических баз данных, которые расположены на разных компьютерах и связаны между собой сетью. Они предназначены для обработки и хранения данных в распределенной среде, где данные могут быть доступны и обрабатываться из разных мест. Данный класс БД обеспечивает высокую доступность и отказоустойчивость данных, а также позволяет распределять нагрузку по разным серверам.

Хранилища данных представляют собой специально организованные базы данных, которые объединяют данные из различных источников и предоставляют единое хранилище для анализа и отчетности. Они используются для хранения больших объемов данных, которые могут быть использованы для принятия стратегических решений и планирования. Такие хранилища обеспечивают высокую производительность при выполнении сложных аналитических запросов и поддерживают механизмы агрегации и индексации данных.

В общем по функциональному назначению БД можно разделить на два вида:

OLAP (англ. Online Analytical Processing, интерактивная аналитическая обработка) – технология обработки данных, заключающаяся в подготовке суммарной (агрегированной) информации на основе больших массивов данных, структурированных по многомерному принципу.

OLTP (англ. Online Transaction Processing, интерактивная обработка транзакций) – обработка транзакций в реальном времени. Способ организации БД, при котором система работает с небольшими по размерам транзакциями, но идущими большим потоком, и при этом клиенту требуется от системы минимальное время отклика.

4. ИСКУССТВЕННЫЕ НЕЙРОННЫЕ СЕТИ

4.1. КЛАССИФИКАЦИЯ МАШИННОГО ОБУЧЕНИЯ

Возможность передавать решение различных задач машинам, что позволит избежать повторяющихся действий, в качестве теоретической возможности постулировалась на протяжении веков и до сегодняшнего дня продолжает будоражить воображение. В настоящее время искусственный интеллект (ИИ) нашел применение во многих областях, включая автоматизированную поддержку клиентов, передовые системы помощи водителю и т.д. Машинное обучение (МО) заимствует свои основы из концепций ранних исследований ИИ, а ее подход к решению проблем завязан на статистических данных. Поэтому машинное обучение отделилось от ИИ и сформировалось в виде отдельной области исследований, поскольку ИИ больше направлен на построение автоматизированных экспертных систем, моделируемых с использованием символического языка. МО сосредоточено на построение моделей из статистических данных и теории вероятностей [39].

Томас Байес заложил основы статистического обучения в своем эссе по теории вероятностей (1763) [40]. Предложенные ключевые концепции, такие как условная вероятность, позже объединились в то, что сейчас называется теоремой Байеса, они имеют огромное значение при формулировании некоторых ранних методов МО, таких как наивные Байесовские или Марковские цепи [38]. Ранние наработки в МО также характеризовались акцентом на символические представления усвоенных знаний, такие как производственные правила, деревья решений и логические формулы [39]. Исследования продолжались, и были сделаны другие открытия, одним из наиболее важных стало изобретение первой машины с нейронной сетью (1951). Однако так продолжалось до тех пор, пока Фрэнк Розенблатт не изобрел персептрон (1957) [27], алгоритм классификации, который делает свои предсказания на основе линейной предикторной функции, объединяющей набор весов с вектором признаков, который начал использоваться в искусственных нейронных сетях (ИНС) и привлекать все больше внимания со стороны других исследователей. В 1986 году был предложен процесс обратного распространения, который положил начало дальнейшему развитию ИНС [25], а в 1990-х гг. ИНС и методы опорных векторов (МОВ) стали широко популярны, поскольку доступная вычислительная мощность начала увеличиваться. С тех пор было сделано много других открытий [5], и сегодня МО – это обширная область, которая включает в себя множество методов [8, 29].

Успех МО обеспечило его использование для создания виртуальных личных помощников, почтовых спам-фильтров, рекомендаций по продуктам или приложений для обнаружения мошенничества, которыми ежедневно пользуются миллиарды людей по всему миру. Несмотря на свои практические и коммерческие успехи, МО остается молодой областью со многими недостаточно изученными исследовательскими возможностями. Некоторые из этих

возможностей можно увидеть, сопоставив современные подходы МО с типами обучения, которые мы наблюдаем в естественных системах, таких как люди и другие животные, организации, экономика и биологическая эволюция.

Системы МО могут быть классифицированы по широким категориям в зависимости от того:

- обучаются ли они под наблюдением человека (контролируемое, неконтролируемое, полуконтролируемое обучение и обучение с подкреплением);
- могут ли они обучаться с новыми данными на лету (онлайн или пакетное обучение);
- сравнивают ли они только новые точки данных с уже известными точками данных или изучают закономерности, присутствующие в данных (обучение на основе экземпляров или на основе моделей).

Эти критерии классификации систем МО не являются исключительными, т.е. их можно комбинировать, и методы МО могут по-разному использоваться в разных системах. Например, ИНС может быть системой пакетного обучения, основанной на моделях, с контролируемым обучением, но модели также могут обучаться онлайн с подкреплением.

4.1.1. Контролируемое и неконтролируемое обучение

Объем и тип контроля, который система получает во время обучения, определяют ее тип в соответствии с этими критериями.

При контролируемом обучении данные, которые передаются в систему, включают желаемое решение, называемое меткой. Задача классификации является типичным примером систем такого типа, поскольку алгоритму задаются точки данных, которые классифицируются с правильной меткой, и он изучает статистические свойства или закономерности в этих точках данных для достижения указанной метки. Задача регрессии также может быть решена с использованием контролируемого обучения системы, поскольку цель состоит в том, чтобы предсказать числовое значение, заданное набором признаков (входных переменных). Для обучения этих систем необходимо снабдить алгоритм точками данных, каждая из которых содержит набор признаков и числовое значение, которое он должен предсказать. Затем система изучает соотношение между выходными данными и набором признаков. Некоторые из наиболее важных алгоритмов обучения с контролем – это линейная регрессия, логистическая регрессия, полиномиальная регрессия, деревья решений, методы опорных векторов (см. рис. 16, *a*) и ИНС.

При неконтролируемом обучении данные, используемые для обучения системы, не помечаются. Эти алгоритмы формируют группы точек данных на основе их характеристик, не присваивая класс каждой группе, поскольку нет меток. Это основное направление алгоритмов кластеризации и визуализации. Уменьшение размерности – задача, связанная с этими системами, и ее цель состоит в упрощении данных при сохранении информации путем объединения

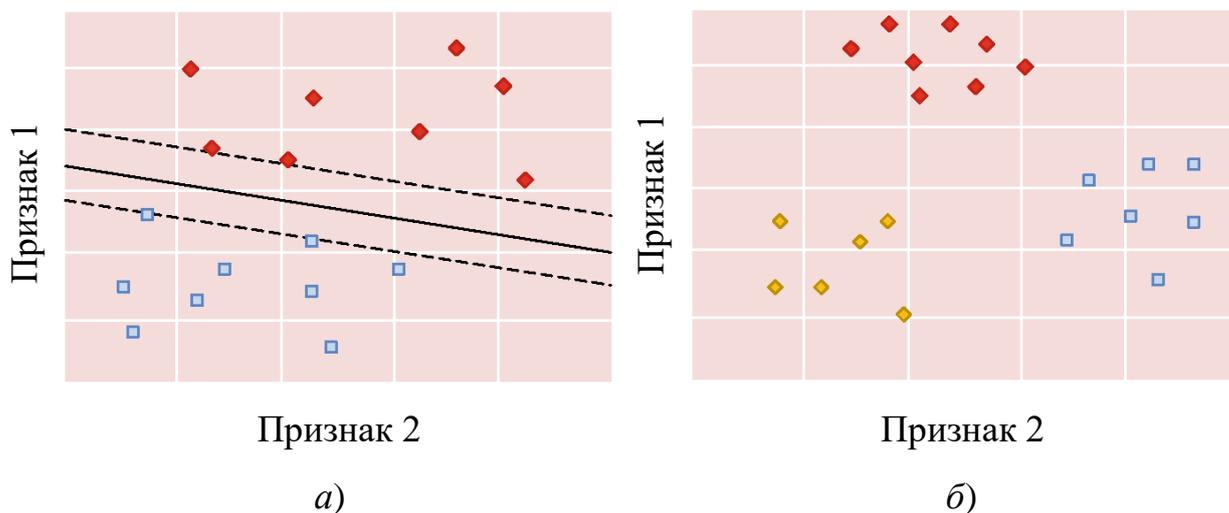


Рис. 16. а) пример алгоритма обучения с контролем, в котором для классификации данных на два разных класса применялся метод опорных векторов; б) пример неконтролируемого алгоритма обучения, в котором использовался метод k -средних для классификации данных на три разных класса

объектов, которые сильно коррелируют. Этот метод называется извлечением объектов, и он используется для уменьшения входных размеров и упрощения визуализации сгруппированных данных точки. Эти системы полезны для обнаружения аномалий, поскольку они группируют «нормальные» точки данных. Если новая точка данных находится далеко от «нормальной» группы, это может быть аномалией. Некоторыми из наиболее важных алгоритмов обучения без контроля являются алгоритмы кластеризации, такие как метод k -средних значений, и алгоритмы визуализации и уменьшения размерности, такие как анализ главных компонент (см. рис. 16, б).

При обучении под наблюдением данные, используемые для обучения системы, частично помечены. Обычно обучающие данные содержат много немаркированных точек данных и несколько помеченных. Большинство алгоритмов обучения с полуконтролем представляют собой комбинации контролируемых и неконтролируемых алгоритмов обучения. Например, сети глубокого убеждения или глубоких сетей доверия основаны на неконтролируемых компонентах, называемых ограниченными методами Больцмана, которые обучаются неконтролируемым образом, а затем вся система настраивается с использованием методов контролируемого обучения.

Обучение с подкреплением подразумевает существование агента, который может наблюдать за окружающей средой и взаимодействовать с ней, выбирая и выполняя действия. Агент получает вознаграждение в зависимости от выбранного им действия (вознаграждение может быть положительным или отрицательным). Вознаграждение необходимо, чтобы узнать, какая стратегия, называемая политикой, является наилучшей, чтобы максимизировать вознаграждение с течением времени. Политика – это функция, которая определяет,

какое действие агент должен выбрать, когда он находится в данной ситуации. Эти системы используются в роботах, чтобы научить их ходить. Обучение с подкреплением также использовалось в программе DeepMind AlphaGo, которая победила чемпиона мира по игре Go [3]. Программа обучила свою политику, проанализировав миллионы игр, а затем сыграв против самой себя.

4.1.2. Пакетное и онлайн-обучение

Эта классификация различает системы по их способности к постепенному обучению на основе потока поступающих данных. При пакетном обучении система обучается с использованием всех доступных данных. Этот процесс занимает много времени и требует значительных вычислительных ресурсов. По этой причине обучение обычно проводится до того момента, когда система будет способна адекватно прогнозировать результаты для новых данных. Если система сначала обучается, а затем используется, то это называется автономным обучением. Чтобы обучить уже обученную систему пакетного обучения новым данным, необходимо обучить новую версию системы с нуля с новым набором данных (старые данные дополнены новыми данными), а затем заменить старую систему на заново обученную. Поскольку выполнение этого процесса может занимать несколько часов, в коммерческих системах эта операция обычно выполняется еженедельно. Плюсом является то, что легко автоматизировать обучение, оценку и запуск новой системы. Еще одним ограничением этих систем является то, что при больших наборах данных может не хватить вычислительных ресурсов для обучения с полным набором данных, что делает невозможным использование алгоритма пакетного обучения.

При онлайн-обучении система обучается поэтапно путем последовательной подачи новых данных, либо единичными данными, либо небольшими группами данных, называемых мини-пакетами. В отличие от систем пакетного обучения, системы онлайн-обучения быстрее и дешевле с точки зрения вычислительных ресурсов, при этом такие системы могут узнавать о новых данных непосредственно во время работы. Эти системы идеальны, когда новые данные доступны непрерывным потоком или когда вычислительные ресурсы ограничены. Недостатком является тот факт, что эти системы постоянно потребляют новые данные и изучают их, поэтому, если при поступлении в систему плохих/неверных данных ее производительность начнет снижаться.

4.1.3. Обучение на основе примеров или моделей

Этот подход применяется в системах, которые используются для обобщения новых данных. Почти каждая задача МО в качестве своей цели имеет предсказания, поэтому, даже если она имеет 100 %-ную точность на обучающем наборе, в конечном итоге ей приходится обобщать данных на примерах, которые она никогда раньше не видела.

Системы обучения на основе примеров используют для обобщения меры сходства с новыми данными. Эта мера сходства используется для сравнения новых данных с уже известными данными и вывода результата на основе этого значения. Алгоритмы кластеризации представляют собой пример таких систем, поскольку новые данные приписываются существующему кластеру на основе расстояний до этих кластеров.

Системы обучения на основе моделей, как следует из названия, строят модель на основе обучающих данных для обобщения с новыми данными. Обычно модели определяются набором параметров, которые необходимо изучить на этапе обучения. Чтобы измерить, какой набор параметров обеспечивает лучшую производительность, определяется функция затрат (полезности), которая оценивает, насколько плохой (хорошей) является модель. ИНС являются ярким примером такого рода систем, поскольку они сами определяют модель.

4.2. ПРОБЛЕМЫ ПРИ СОЗДАНИИ СИСТЕМ МАШИННОГО ОБУЧЕНИЯ

Система машинного обучения – это обучающий алгоритм, тренируемый на данных. При проектировании такой системы могут возникнуть две проблемы: проблема с алгоритмом или проблема с данными, используемыми для его обучения. Большинству алгоритмов МО для правильной работы требуется большой объем данных, и даже для простых задач набор данных должен содержать тысячи примеров, в то время как для более сложных программ, таких как распознавание изображений, примеров должно быть миллионы. При наличии достаточного количества данных различные алгоритмы МО могут одинаково хорошо работать со сложной задачей. Это было показано Банко и Брилл в [4], где разные алгоритмы показали одинаково хорошую производительность в задаче устранения неоднозначности естественного языка. Эти результаты предполагают, что, возможно, важнее потратить время и деньги на получение большего количества и более качественных данных, чем на разработку алгоритма. Однако наборы данных малого и среднего размера по-прежнему очень распространены так как получить больше данных непросто и недешево.

Чтобы хорошо обобщать, алгоритм должен иметь доступ ко всем случаям, которые он должен обрабатывать во время обучения; то есть обучающие данные должны содержать примеры каждого случая, о котором алгоритму необходимо узнать. Если это не так, система может начать демонстрировать низкую точность, когда ей начнут поступать входные данные, не представленные в обучающих данных. Кроме того, обучающие данные могут содержать ошибки, выбросы и шум.

Это затрудняет для системы изучение закономерностей в данных, что влияет на ее производительность. Часто стоит очистить данные, прежде чем вводить их в алгоритм обучения.

Более того, система может обучаться только в том случае, если данные содержат достаточно релевантных характеристик и несколько нерелевантных. Таким образом, важно выполнить процесс, называемый инженерией признаков, который включает в себя выбор наиболее полезных признаков, объединение существующих признаков в один, более полезный (например, с использованием алгоритмов уменьшения размерности) и создание новых функций путем получения новых данных или расширения имеющихся данных (например, полиномиальные функции, сдвиг изображений и т.д.). Это будет гарантировать, что система обучается полезными данными.

Предотвращение переобучения является основной задачей при проектировании систем МО. Модель переобучения хорошо работает с обучающими данными, но плохо обобщается на новых данных, что является основной целью систем МО: предсказать правильный вывод для данных, которые она ранее не видела. Сложные модели более подвержены переобучению, поскольку им легче обнаруживать закономерности в данных. Однако, если примеров не слишком много или они зашумлены, модель может начать обнаруживать закономерности в самом шуме. Переобучение обычно происходит, когда модель слишком сложна для данного объема данных и/или из-за высокой зашумленности обучающих данных. Чтобы предотвратить переобучение, можно упростить модель, выбрав модель с меньшим количеством параметров, уменьшив количество атрибутов в обучающих данных или ограничив модель. Очевидно, что увеличение количества примеров или уменьшение количества шума (путем удаления выбросов и исправления ошибок) в обучающих данных также помогает предотвратить переобучение.

В отличие от переобучения, недостаточное соответствие также может быть проблемой в системах МО. Недостаточное соответствие может возникнуть, когда модель слишком проста, и поэтому ей не удастся изучить закономерности в данных. Этого можно избежать, сделав модель более сложной, увеличив количество параметров или добавив в процесс обучения улучшенных функций.

4.3. ПЯТЬ ПОДХОДОВ К МАШИННОМУ ОБУЧЕНИЮ

Хотя методы, представленные в этой главе, в основном рассматривают глубокое обучение и ИНС, в этом разделе представлен краткий обзор более широкого спектра алгоритмов МО.

В своей книге *The Master Algorithm* [40] Домингос разделяет все алгоритмы МО в соответствии с пятью различными подходами, основанными на определенных базовых методах и философии. Эти подходы определяются как символисты, байесовцы, коннекционисты, эволюционисты и аналогизаторы. В таблице 1 приведена краткая информация о преимуществах и недостатках каждого принципа.

1. Преимущества и недостатки подходов машинного обучения

Подходы	Репрезентативный алгоритм	Преимущества	Недостатки
Символисты	Обратная дедукция Правила и деревья решений графы знаний	Просты для понимания. Могут обрабатывать как числовые, так и категориальные данные, а также задачи с несколькими выводами	Непрактичный подход, если необходимо принять несколько решений. Слишком сложные деревья могут в конечном итоге плохо обобщать данные, что приведет к переобучению
Байесовцы	Наивные цепи Байеса и Маркова Вероятностный вывод	Для эффективности требуется меньше данных. Быстрое предсказание класса объекта. Нечувствительность к несущественным функциям	Бесполезен при работе с реальными ситуациями, поскольку функции обычно взаимозависимы
Коннекционисты	Искусственные нейронные сети	Подходит для сложных и зашумленных наборов данных. Хорошее обобщение правил между входными объектами и выходными значениями	Требуется повторить алгоритм несколько раз, чтобы получить положительные результаты. Существует вероятность, что произойдет переобучение
Эволюционеры	Эволюционные алгоритмы и генетическое программирование	Подходит для задач с широким диапазоном параметров	Дорогое время вычислений. Может не найти глобальные максимумы
Аналогизаторы	Методы опорных векторов	Может решать как линейные (классификация), так и нелинейные (регрессия) задачи. Может генерировать уникальные решения	Не подходит для сложных и зашумленных наборов данных. Нежелательный выбор при решении проблемы с многомерными данными

Все пять подходов могут предложить что-то уникальное. Главное – понять, какой тип данных доступен, и выяснить, какая методология наилучшим образом подойдет для рассматриваемой проблемы. Разные проблемы требуют разных подходов, и в следующих разделах рассматриваются сильные и слабые

стороны нескольких подходов МО и анализируется их актуальность в контексте автоматизации различных процессов. Некоторые проблемы могут быть решены с использованием только одного метода, в то время как другие могут потребовать сочетания различных методов.

Размер и структура данных являются важными свойствами для раннего анализа. Оценка типа решаемой учебной задачи (контролируемой или неконтролируемой) также является простым способом исключить некоторые варианты. Время вычислений также является важным свойством, которое следует принимать во внимание.

Символисты утверждают, что знание может быть сведено к манипулированию символами, точно также, как математик решает уравнения, заменяя выражения другими выражениями. Вместо того, чтобы начинать с обработки исходных данных и искать обобщенные заключения, обратная дедукция предлагает начать обработку данных с некоторых промежуточных данных и заранее определенных выводов и, по сути, работает в обратном направлении, чтобы заполнить пробелы. Это делается путем выявления недостающих правил, которые соответствуют заранее установленным выводам (очень похоже на решение головоломки). Репрезентативными алгоритмами являются правила и деревья решений.

Символисты, вероятно, представляют собой самый простой для понимания, интерпретации и визуализации набор методов, с которыми можно столкнуться. Правила и дерево решений дают четкое представление о данных и на их основе легко экстраполировать выводы. Они могут обрабатывать как числовые, так и категориальные данные, а также задачи с несколькими выводами. Нелинейности в данных не влияют на производительность деревьев. Тем не менее, сложные наборы данных, где необходимо вывести несколько правил, могут быть не самыми подходящими для этих методов. Сверхсложные деревья могут в конечном итоге плохо обобщать данные, что приводит к переобучению. Некоторые классы также доминируют над другими, если входной набор данных не сбалансирован, поэтому рекомендуется иметь одинаковое количество примеров для каждого класса (это верно почти для всех алгоритмов).

С точки зрения достижения удовлетворительных результатов, глобальный максимум может быть не найден простым запуском одного дерева решений. Более благоприятный набор результатов может быть получен путем обучения нескольких деревьев, где объекты и выборки выбираются случайным образом с заменой. Случайные леса обобщают эту концепцию и используются для преодоления проблемы переобучения, присущей деревьям решений. Как обучение, так и прогнозирование выполняются очень быстро из-за простоты лежащих в их основе деревьев решений. Кроме того, обе задачи могут быть распараллелены простым способом, поскольку отдельные деревья являются полностью независимыми объектами.

Байесовцы прежде всего озабочены неопределенностью. Этот тип обучения оценивает вероятность того, что гипотеза окажется верной, учитывая

априорные знания. Различные гипотезы сравниваются путем оценки того, какие результаты с большей вероятностью произойдут. Это называется вероятностным выводом. Излюбленными алгоритмами этого подхода являются наивные байесовские или Марковские цепи.

Наивный алгоритм Байеса обеспечивает быстрое, масштабируемое построение модели и оценку. Алгоритм может использоваться как для задач бинарной, так и для многоклассовой классификации. Эти модели относительно просты для понимания и построения. Они легко обучаются и не требуют больших наборов данных для получения эффективных результатов. Они также нечувствительны к нерелевантным признакам. Эти преимущества означают, что использование наивного байесовского классификатора часто является хорошим выбором для начальной базовой классификации. Если алгоритм работает надлежащим образом, то возможно получить очень быстрый и интерпретируемый классификатор для решения задачи без особых усилий. Если он работает плохо, следует изучить другие модели. Поскольку этот алгоритм всегда предполагает, что функции независимы, что неверно для большинства реальных ситуаций, во многих случаях эта модель не будет работать должным образом.

Для аналогизаторов ключом к обучению является распознавание сходства между ситуациями и, таким образом, вывод о других сходствах. Обучение сводится к построению аналогий между доступными данными. Подход аналогизаторов более известен благодаря методу опорных векторов. Это один из наиболее эффективных алгоритмов МО и в основном используется для распознавания образов [38]. Он имеет широкий спектр применений, таких как распознавание речи, распознавание лиц и изображений. Это очень мощный алгоритм контролируемого обучения для разделения данных, который строит модель, которая предвидит категорию нового примера, на основе заданного набора рекомендуемых примеров. Он работает по принципу подгонки границы к области точек, которые соответствуют одним и тем же критериям классификации, т.е. принадлежат к одному классу. Как только граница установлена на точках обучающей выборки, для любых новых точек, которые необходимо классифицировать, разработчик должен только проверить, находятся ли они внутри границы или нет. Преимущество МОВ в том, что после установления границы большая часть обучающих данных становится избыточной. Все, что ему нужно – это базовый набор точек, которые могут помочь определить и установить границу. Эти точки данных называются опорными векторами, потому что они «поддерживают» границу. Типы данных для этого алгоритма включают линейные и нелинейные шаблоны. Линейные шаблоны легко различимы и могут быть разделены при малых размерах, но этого нельзя сказать о нелинейных шаблонах.

Последним необходимо манипулировать, чтобы данные стали разделимыми, например, с помощью функций ядра. Еще одно преимущество МОВ заключается в том, что они очень хорошо обобщают новые выборки данных. Когда достигается оптимальный набор гиперплоскостей, разделяющих данные,

МОВ может выдавать уникальные решения, что является фундаментальным отличием этого метода от ИНС. Последний дает множество решений, основанных на локальных минимумах, которые могут быть неточными для разных тестовых данных. С точки зрения недостатков, МОВ может оказаться не самым желательным выбором при решении проблем с многомерными данными. Кроме того, они в значительной степени зависят от выбора ядра и его параметров, и даже в этом случае полученные результаты могут быть недостаточно прозрачными, чтобы экстраполировать какие-либо значимые выводы.

Эволюционеры считают, что основой всего обучения является естественный отбор. По сути, эволюционный алгоритм — это метаэвристический алгоритм оптимизации, используемый в ИИ, который использует механизмы, основанные на биологической эволюции. Помимо своей способности искать множество решений в больших пространствах данных, эти алгоритмы способны поддерживать разнообразную совокупность решений и использовать сходство решений с помощью механизмов рекомбинации, мутации, размножения и отбора [25]. Функция соответствия отвечает за анализ качества предлагаемых решений, вывод значений, которые затем будут сравниваться с предопределенной функцией затрат или целевой функцией, или набором оптимальных компромиссных значений в случае двух или более противоречащих друг другу целей. Эволюционные алгоритмы представляют собой набор современных приемов, успешно используемых во многих приложениях большой сложности. Самое непосредственное преимущество эволюционных вычислений заключается в том, что они концептуально просты. Практически любая проблема, которая может быть сформулирована как задача оптимизации функции, может быть смоделирована с помощью эволюционного алгоритма [27]. Эти алгоритмы способны к самооптимизации, и их процессы в высокой степени параллельны. Это означает, что оценка каждого полученного решения может выполняться параллельно. Только механизм отбора требует некоторой последовательной обработки. Эволюционные алгоритмы также могут использоваться для адаптации решений к изменяющимся условиям. Обычно нет необходимости получать наборы данных случайным образом и перезапускать модель, когда требуется получить новые решения. Это связано с высокой гибкостью алгоритма, что делает доступные наборы данных прочной основой для дальнейших улучшений. К недостаткам этого алгоритма относятся: отсутствие гарантии нахождения глобальных максимумов и большое время вычислений.

Для коннекционистов обучение — это то, что делает мозг, и поэтому цель состоит в его обратном проектировании. Модели, принадлежащие к этому классу, пытаются построить сложные сети, состоящие из узлов, напоминающих нейроны головного мозга, и регулировать силу каждого соединения, сравнивая полученные выходные данные с желаемыми. Их излюбленным алгоритмом является обратное распространение, метод, обычно применяемый к ИНС.

Коннекционисты предлагают широкий спектр приложений. За последние несколько лет глубокое обучение стало довольно популярным в обработке

изображений, распознавании речи и других областях, где доступен большой объем данных. ИНС, наиболее распространенная реализация подхода, может дать некоторые впечатляющие результаты, поскольку они могут обобщать правила между входными признаками и выходными переменными, учитывая, что в исходном наборе данных достаточно примеров.

Этот метод хорошо подходит для задач, в которых обучающие данные зашумлены и работе со сложными выходными данным, например, данными с камер и микрофонов. Выходные данные целевой функции могут быть дискретными, вещественными или вектором из нескольких вещественных или дискретных атрибутов. Это подходящий метод для наборов данных, содержащих обучающие примеры с ошибками, и, хотя обучение ИНС может занять относительно много времени, оценка обученной сети обычно выполняется очень быстро. Какими бы мощными ни были эти методы, достижение наиболее благоприятной архитектуры модели является сложным и итеративным процессом. Есть несколько деталей, которые разработчик должен учитывать при разработке ИНС, в которых минимизированы ошибки обучения сети, избегая при этом переобучения, а именно количество слоев, функций активации, функции потерь и используемых оптимизаторов или нормализаторов данных.

4.4. ОБЗОР НЕЙРОННЫХ СЕТЕЙ

Строительным блоком ИНС является искусственный нейрон (также называемый узлом). Операции, которые выполняет каждый нейрон, просты: он суммирует свои взвешенные входные данные, передает полученные результаты через функцию, называемую функцией активации (обычно нелинейную), и выводит это значение, как показано на рис. 17.

Хотя эти блоки выполняют простые операции, когда они сгруппированы в плотные сети, они могут образовывать очень сложные функции [23]. Параметры, определяющие эти функции, являются весами каждого нейрона и определяются во время обучения.

Фундаментальной структурой для искусственной нейронной сети является многослойный персептрон (МП), показанный на рис. 18.

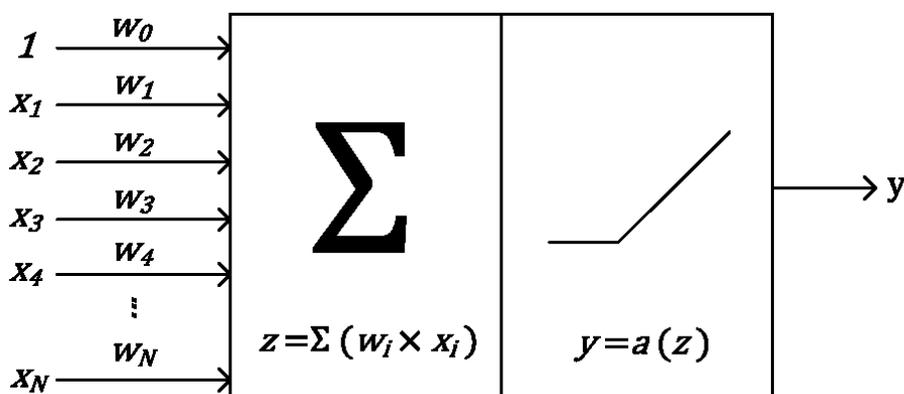


Рис. 17. Представление искусственного нейрона

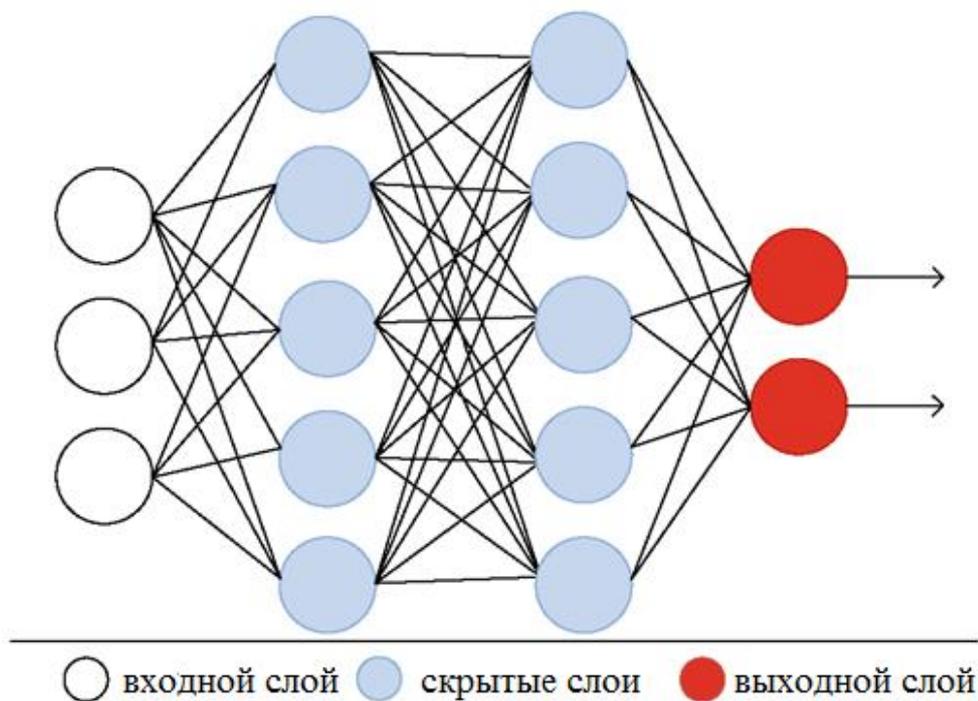


Рис. 18. Полносвязная ИНС с двумя скрытыми слоями

Он состоит из трех или более слоев узлов (или нейронов): одного входного слоя, одного выходного слоя и одного или более скрытых слоев. Входной уровень отвечает только за распределение входных данных на следующий уровень: он не выполняет с ними никаких операций. Количество нейронов во входном слое всегда равно количеству входных данных модели. Нейроны в скрытых слоях принимают выходные данные нейронов предыдущего слоя и выводят их на следующий слой. Эти слои могут содержать любое количество нейронов. Выходной слой получает выходные данные от последнего скрытого слоя, и его выходные данные соответствуют выходным данным модели. Также возможно спроектировать сеть только с входным и выходным слоями (без скрытых слоев). В наиболее распространенной слоистой структуре в более ранних моделях выход каждого нейрона был бы входом для каждого нейрона в следующем слое, как это представлено в ИНС на рис. 18. Эти слои называются полностью подключенными слоями, возможны и другие конфигурации, такие как сверточные слои, где выходные данные узлов подключены только к некоторым узлам на следующем уровне [25], или остаточные слои, где выходные данные некоторого слоя передаются на следующий уровень и добавляются перед функцией активации следующего слоя [5].

Многослойный перцептрон – не единственный тип ИНС, существуют и другие модели, такие как сверточные нейронные сети (СНС) и рекуррентные нейронные сети (РНС), которые хорошо подходят для решения конкретных типов задач. Например, СНС чрезвычайно хорошо подходят для задач обработки изображений, где в дополнение к полностью связанным слоям в них есть сверточные слои, соединенные с узлами, которые преобразуют фильтры во

входную матрицу для извлечения признаков. Веса фильтров определяются во время обучения. РНС структурированы таким образом, что позволяют обрабатывать последовательности переменной длины. Они хорошо подходят для задач обработки естественного языка, например, для распознавания рукописного ввода и речи, поскольку часть выходных данных в момент времени t передается в качестве входных данных для времени $t + 1$ (и $t - 1$ в двунаправленных сетях), что представляет собой способ запоминания, который в случае задач естественного языка вводит чувствительность к контексту.

4.4.1. Размер модели

Гиперпараметрами, определяющими количество узлов ИНС, являются количество слоев (или глубина) сети и количество узлов в каждом слое (или ширина). Кроме того, структура соединения (полное соединение, общие веса и т.д.) также влияет на общее количество параметров, которые должны быть установлены во время обучения. Количество узлов в ИНС является важным гиперпараметром и определяет сложность функции, которую может представлять ИНС. В целом, большее количество уровней (более глубокая модель) приведет к тому, что ИНС сможет изучать более сложные функции при том же количестве узлов [8], но это также увеличивает сложность обучения сети.

Глубокие сети трудно обучать из-за числовой нестабильности, например, проблемы исчезающего или взрывного градиента, могут значительно ухудшить способность эффективно обновлять веса.

Что касается количества нейронов, то установить количество единиц во входном и выходном слое просто, поскольку они представляют собой количество входных и выходных переменных соответственно.

Однако скрытые слои могут создавать проблему, когда слишком много нейронов представляют более сложные функции, но отрицательно влияют на время обучения модели, поскольку количество вычисляемых весов значительно увеличивается. Объективного способа определить оптимальное количество нейронов в каждом скрытом слое ИНС не существует.

Некоторые приемы приводят к хорошим результатам в определенных задачах, но они плохо обобщаются, что в конечном счете превращает процесс выбора количества нейронов в скрытых слоях в процесс проб и ошибок.

Чтобы иметь отправную точку для определения количества нейронов (N_h для скрытого слоя h), можно воспользоваться некоторыми приемами, с которыми можно поэкспериментировать:

- для сетей с двумя скрытыми уровнями, формулы

$$N_{h=1} = \sqrt{(m + 2)N} + 2 \sqrt{\frac{N}{m + 2}}$$

и

$$N_{h=2} = m \sqrt{\frac{N}{m + 2}},$$

где N – количество выборок, которые должны быть изучены с низкой ошибкой, а m – количество выходных переменных, что приведет к созданию сети, которая может изучить N выборок (пар ввода-вывода) с небольшой ошибкой [29];

– формула $N_{h=L} = \frac{N_{in} + \sqrt{N_p}}{L}$ была протестирована на 40 различных тестовых примерах. Исходя из полученных результатов, авторы заявили, что он может быть использован для вычисления оптимального количества нейронов в скрытых слоях. В этой формуле N_{in} – количество входных нейронов, N_p – количество входных выборок, а L – номер скрытого слоя [3];

– Гласснер Э. предлагает использовать теорему для вычисления необходимого количества скрытых узлов в слоях, чтобы заставить ИНС достичь заданного порядка аппроксимации по отношению к функции, которую она пытается смоделировать. Порядок аппроксимации N подразумевает, что производная порядка N в начале координат одинакова для функции $g(x)$, определяемой ИНС, и для функции, которую она пытается смоделировать, $f(x)$, или: $g^{(N)}(0) = f^{(N)}(0)$. Для достижения порядка аппроксимации N количество скрытых единиц узлов во всех скрытых слоях должно соответствовать условию (1), где n_0 – количество входных данных модели [4]

$$\sum_{h=1}^L N_h \geq \frac{\binom{N + n_0}{n_0}}{n_0 + 2},$$

где

$$\binom{N + n_0}{n_0} \leq n_0^2 + 3n_0 + 2\sqrt{n_0^3} + 4\sqrt{n_0}; \quad (1)$$

отсюда

$$\sum_{h=1}^L N_h \geq 2\sqrt{\binom{N + n_0}{n_0}} + 2n_0 + 2 - n_0 - 3.$$

Существует некоторое эмпирическое правило, которое можно использовать для выбора общего количества скрытых нейронов, например:

– количество скрытых нейронов должно находиться между количеством нейронов входного и количеством нейронов выходного слоя;

– количество скрытых нейронов должно составлять 2/3 от количества нейронов входного слоя плюс количество нейронов выходного слоя;

– количество скрытых нейронов должно быть менее чем в два раза больше количества нейронов входного слоя.

– одно из простейших правил – метод проб и ошибок. Он работает путем обучения сетей с различными комбинациями количества единиц для скрытых слоев и сохраняет ту, которая дает лучшие результаты. Проблема этого метода в том, что он отнимает много времени, поскольку обучение ИНС обычно занимает много времени. Первоначально количество скрытых нейронов может быть задано с помощью одной из приведенных выше эвристик, но, в конце концов, тонкая настройка этих параметров должна быть

скорректирована с использованием этого метода в поисках модели, которая приводит к лучшим результатам.

Как видно, установка соответствующей глубины и количества узлов имеет первостепенное значение, поскольку это определяет выразительность моделей. Однако трудно угадать правильное количество слоев и количество узлов в каждом слое, которые сформируют наилучшую модель. Следовательно, чтобы найти наилучшую форму ИНС для решения желаемой задачи, хорошим первым шагом является обучение сети с двумя слоями, что соответствует простой многомерной линейной модели. Оценка этой модели позволяет нам оценить насколько сложны данные и обеспечивает основу для улучшения, затем увеличиваем размер модели (ширину и глубину), пока не получим приемлемую производительность.

После этого можно настроить другие гиперпараметры для дальнейшего повышения точности ИНС при сохранении общей структуры.

4.4.2. Функции активации

На заре появления ИНС наиболее распространенными функциями активации были тождественность, сигмоид и гиперболический тангенс, показанные на рис. 19.

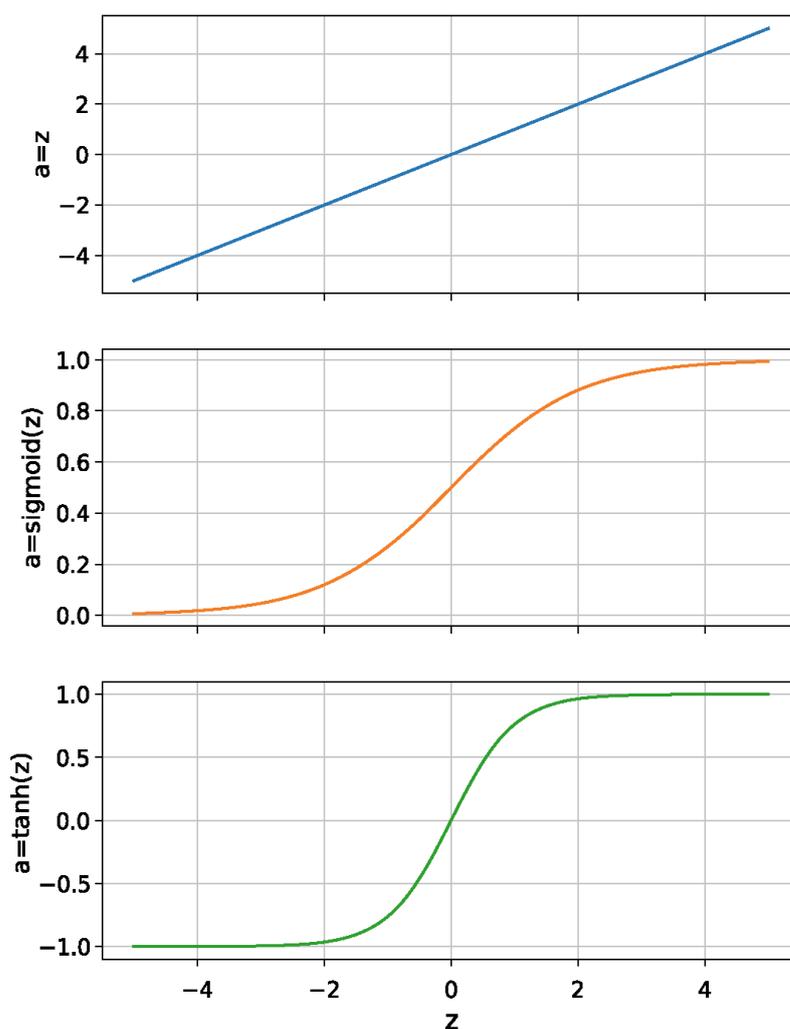


Рис. 19. Линейная, сигмовидная и функция гиперболического тангенса

Тождественность используется только на выходном слое в задачах регрессии, так что на выходе модели может быть любое значение (положительное или нет). Однако в скрытых слоях возможно использование нелинейных функции, потому что такие функции легко представить с помощью ИНС, а сигмоида и гиперболический тангенс начали использоваться в качестве функций активации, поскольку они реализуют поведение биологических нейронов [29].

Однако в 2010 г. Глорот и Бенгио [39] пришли к выводу, что проблема с исчезающим/взрывающимся градиентом была вызвана неправильным выбором функции активации. После этого открытия было предложено новое семейство функций активации. Все началось с выпрямленной линейной функции активации (ReLU) [40], как описано в (2)

$$\text{ReLU}(z) = \max(0, z) . \quad (2)$$

ReLU лучше работает в глубоких нейронных сетях, поскольку она не насыщается при положительных значениях; кроме того, и функцию, и градиент очень легко вычислить. Тем не менее, как только сумма взвешенных входных данных нейрона становится отрицательной, нейрон выводит 0 с градиентом, также равным нулю, что может легко привести к среднему градиенту, равному нулю, как только это произойдет, веса перестают обновляться, что приводит к проблеме, известной как умирающий ReLU. Для решения этой проблемы было предложено несколько вариантов ReLU. Дырявый ReLU (lReLU) определяется в соответствии с (3)

$$\text{lReLU}(z) = \max(\alpha z, z) . \quad (3)$$

Параметр α управляет наклоном (утечкой) функции при $z < 0$, и его значение обычно близко к 0. $\alpha = 0,01$ считается небольшой утечкой, а $\alpha = 0,2$ считается огромной утечкой. Рандомизированный дырявый ReLU действует как дырявый ReLU, но параметр α выбирается случайным образом в заданном диапазоне во время обучения и устанавливается на среднее значение во время тестирования. Параметрический дырявый ReLU превращает параметр α в параметр, который может быть изучен во время обучения с использованием метода обратного распространения. Было показано, что дырявый ReLU и его варианты превосходят функцию стандартного ReLU для сверточных нейронных сетей [38].

В 2015 году Клеверт и другие [27] предложили новую функцию активации под названием экспоненциальный линейный агрегат (ELU), которая превзошла все варианты ReLU, сократив время обучения и добившись лучшей производительности в тестируемых задачах. Функция ELU определяется по формуле (4). Гиперпараметр α определяет значение, к которому стремится функция, когда z становится очень большим отрицательным числом

$$ELU(z) = \begin{cases} \alpha(e^z - 1) & \text{if } z < 0; \\ z & \text{if } z \geq 0. \end{cases} \quad (4)$$

ELU похож на дырявый ReLU, но он более гладкий для всех значений z (в том числе около $z = 0$), что помогает ускорить алгоритмы обучения. Несмотря на то, что он в конечном итоге насыщается при $z = 0$, он по-прежнему уменьшает проблему возникновения исчезающих градиентов и имеет ненулевой градиент в отрицательном спектре, избегая проблемы умирающего ReLU. Основным недостатком использования этой функции активации является то, что она вычисляется медленнее, чем варианты ReLU, что замедляет ее работу во время тестирования, даже несмотря на то, что во время обучения она сходится быстрее. На рисунке 20 показаны функции ReLU, lReLU и ELU.

4.4.3. Как искусственные нейронные сети «учатся»

Процесс обучения ИНС соответствует решению задачи оптимизации поиска параметров модели, которые минимизируют некоторую меру погрешности модели, J . Двумя наиболее распространенными функциями погрешности являются среднеквадратическая ошибка (СКО) и средняя абсолютная ошибка (САО), как определено в (5), где M – количество выборок, x_i – входной вектор, y_i – ожидаемый результат, а w – веса модели

$$MSE(x, y, w) = \frac{1}{M} \sum_{i=1}^M (ANN(x_i, w) - y_i)^2;$$

$$MAE(x, y, w) = \frac{1}{M} \sum_{i=1}^M |ANN(x_i, w) - y_i|. \quad (5)$$

Методы, используемые для решения этой задачи оптимизации, основаны на градиентном спуске, итеративном процессе, при котором функция потерь перемещается в направлении, противоположном градиенту. В случае обучения ИНС градиент функции потерь эффективно вычисляется с использованием обратного распространения ошибки [25]. Во время обучения производная ошибки по отношению к каждому весу передается в обратном направлении по сети (отсюда и название метода обратного распространения) и используется для обновления весов в соответствии с (6)

$$w^{(n+1)} = w^{(n)} - \eta \frac{\partial J}{\partial w} (w^{(n)}), \quad (6)$$

где J – функция потерь; w – общий вес, $w^{(n+1)}$ – значение веса при его обновлении, а η – скаляр, называемый шагом обучения.

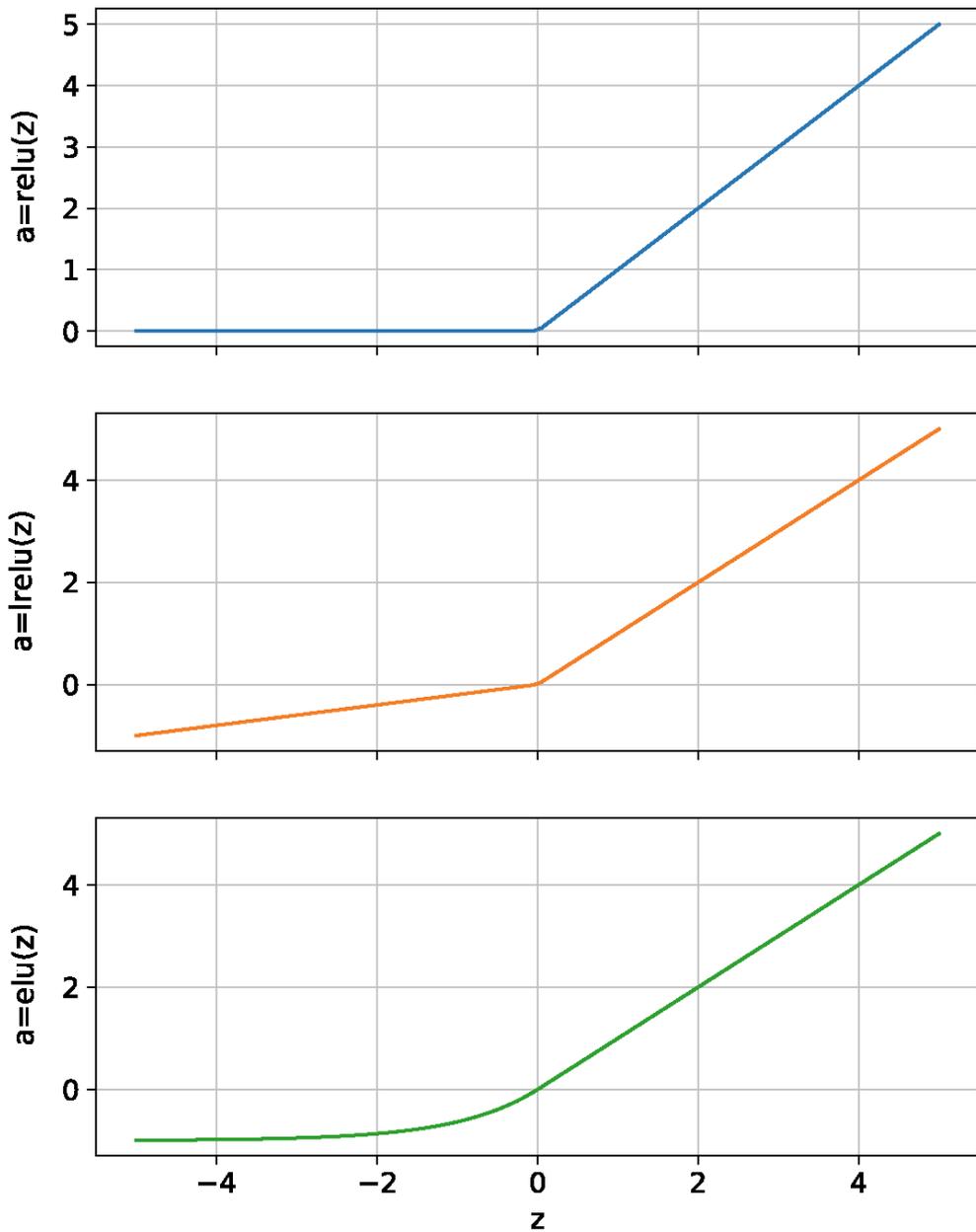


Рис. 20. ReLU, дырявый ReLU при $\alpha = 0,2$, функция ELU для $\alpha = 1$

Обратное распространение использует правило цепочки и представляет простой способ рекурсивного вычисления градиента весовой функции с учетом производных функций активации. Правило обратного распространения представлено в (7), где u_w – это входные данные сети; w – веса в сети прямого распространения, а v_w – входные данные сети этого веса в сети обратного распространения

$$\frac{\partial J}{\partial w}(w^{(n)}) = u_w v_w . \tag{7}$$

Скорость обучения контролирует влияние градиента на обновление весов. Большая скорость обучения может привести к тому, что алгоритм выйдет

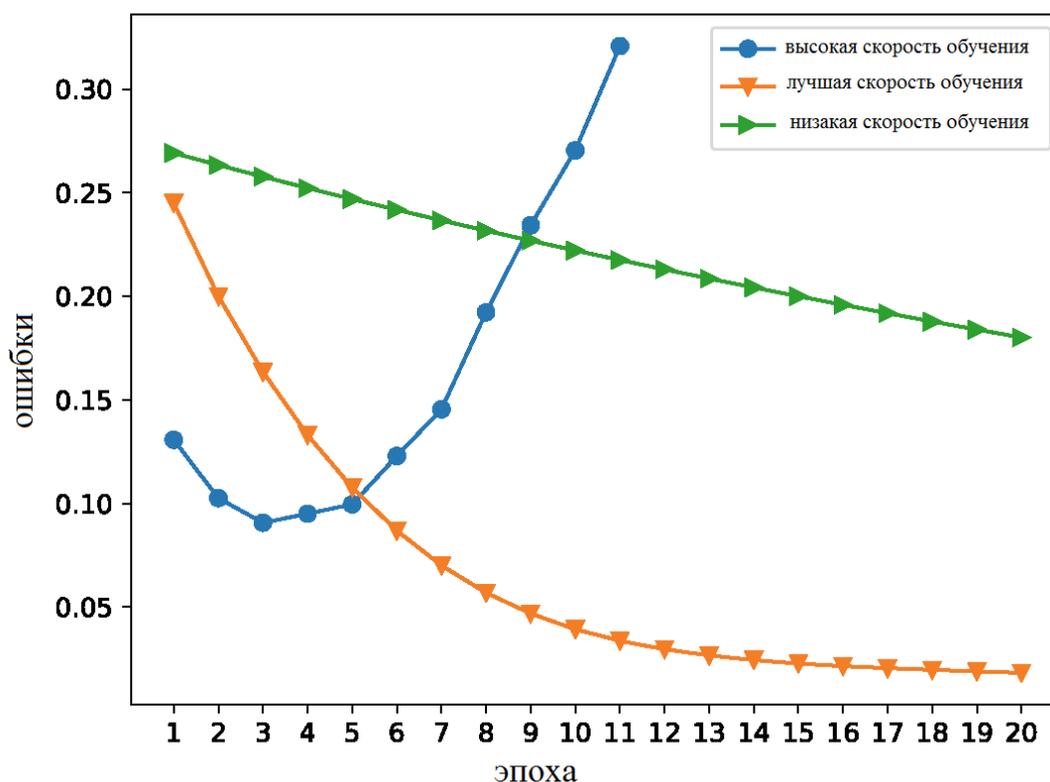


Рис. 21. Эволюция ошибки при различных скоростях обучения

за пределы минимума функции потерь или, в худшем случае, отклонится от минимума. Низкая скорость обучения делает конвергенцию слишком медленной, как показано на рис. 21. Скорость обучения является важным гиперпараметром, поскольку она управляет сходимостью или расхождением алгоритма оптимизации. Однако существует широкий диапазон значений этого параметра, который гарантирует, что алгоритм оптимизации найдет минимум функции потерь. Обычная процедура заключается в том, чтобы сначала попробовать увеличить скорость обучения (0,1 или 0,01), а затем начать итеративно уменьшать величину этого значения, если результаты улучшаются.

Поскольку количество выборок в данных M становится больше, непрактично вычислять весовую функцию по всему набору данных. Кроме того, за счет более частого обновления весов (с меньшим количеством вычислений за каждый раз) можно эффективно найти минимум. Именно так обучаются современные ИНС, используя стохастический градиентный спуск (СГС) [5, 8], где веса обновляются после отправки мини-пакета примеров в сеть. Размер пакета влияет не только на время выполнения, но и на свойства минимумов. Большой размер мини-пакета приводит к сходимости алгоритма к резким минимумам, которые имеют тенденцию к плохому обобщению, в то время как небольшие мини-пакеты имеют тенденцию к плоским минимизаторам, которые имеют тенденцию к лучшему обобщению [29]. Считается, что большой

размер пакета составляет около 10% от количества примеров в наборе данных, а маленький – около 32, поэтому число между этими пределами должно приводить к приемлемым результатам при сохранении разумного времени обучения. Вместо того, чтобы контролировать количество обновлений весов, при обучении ИНС обычно измеряют длину повторений в количестве эпох, то есть количество раз, когда алгоритму представляются все данные обучения.

4.4.4. Оптимизаторы

Чтобы ускорить конвергенцию процесса обучения, т.е. уменьшить количество итераций, необходимых для достижения минимума функции потерь, были предложены различные варианты градиентного спуска. Для сравнения обновления веса w , ранее представленное в (6) с использованием оптимизации градиентного спуска, (6) переписано как (8). Член ∂J по-прежнему вычисляется путем обратного распространения. Изменяется только способ вычисления слагаемого Δw

$$w^{(n+1)} = w^{(n)} - \Delta^{(n+1)}w, \text{ где } \Delta^{(n+1)}w = \eta \frac{\partial J}{\partial w}(w^{(n)}). \quad (8)$$

Алгоритм передачи СНО импульса был предложен Борисом Поляком в 1964 году [3], и его целью является имитация ощущения ускорения обучения. Алгоритм учитывает прошлые градиенты, и, если градиент продолжает двигаться в одном и том же направлении в течение нескольких итераций (одно и то же направление означает один и тот же сигнал), обновление веса ускоряется по мере выполнения итерации, что облегчает выход из плато в функции потерь. Когда градиент меняет направление, алгоритм уменьшает пройденный шаг. Обновление веса w с использованием импульса задается по формуле (9). В нем вводится новый гиперпараметр β . Считается, что значение 0,9 позволяет достичь хороших результатов

$$\Delta^{(n+1)}w = \beta \Delta^{(n)}w + \eta \frac{\partial J}{\partial w}(w^{(n)}). \quad (9)$$

В 1983 году Юрий Нестеров [4] предложил вариант оптимизации импульса, который почти всегда работает лучше, чем обычная версия. Идея ускоренного градиента Нестерова (УГН) заключается в измерении градиента функции потерь не в локальном положении, а немного впереди в направлении импульса, как показано в (10)

$$\Delta^{(n+1)}w = \beta \Delta^{(n)}w + \eta \frac{\partial J}{\partial w}(w^{(n)} + \beta \Delta^{(n)}w). \quad (10)$$

Эта настройка работает, поскольку направление импульса обычно указывает на минимум функции потерь, а это означает, что она ускорит процесс оптимизации, поскольку градиент будет рассчитываться ближе к оптимальному значению.

Градиентный спуск начинает продвигаться в направлении самого крутого склона, и когда он достигает конца этого склона, он начинает двигаться в направлении следующего склона.

Адаптивный градиент (АдГрад) был предложен в [39], и его цель – заранее обнаружить изменение направления к оптимальному значению. Это достигается путем уменьшения градиента по самому резкому изменению значений в соответствии с (11), где масштабирующие коэффициенты вычисляются с использованием (12). Поскольку градиент при резком измерении вычисляемых значений изменяется больше, слагаемые становятся больше, что уменьшает шаг. Это делается для того, чтобы алгоритм раньше начал двигаться в направлении оптимального значения. ε – это сглаживающий параметр, позволяющий избежать деления на ноль, и обычно он устанавливается равным 10^{-10} . Преимущество этого подхода заключается в том, что он требует меньшей настройки скорости обучения. АдГрад хорошо работает для более простых квадратичных задач, но может при расчетах перестать приближаться минимумам функции потерь в сложных задачах, таких как ИНС

$$\Delta^{(n+1)}_w = \frac{\eta}{\sqrt{S^{(n+1)} + \varepsilon}} \frac{\partial J}{\partial w} (w^{(n)}), \quad (11)$$

где

$$S^{(n+1)} = \sum_{t=1}^n \left[\frac{\partial J}{\partial w} (w^{(t)}) \right]^2 = S^{(n)} + \left[\frac{\partial J}{\partial w} (w^{(n)}) \right]^2. \quad (12)$$

Среднеквадратичное значение (СК) пытается решить проблему слишком ранней остановки приближения к минимуму функции потерь АдГрад, учитывая только градиенты последних итераций, а не все градиенты с начала выполнения алгоритма. Он использует экспоненциальное затухание для вычисления коэффициентов масштабирования, как показано в (13). Скорость затухания β – это новый гиперпараметр, который необходимо настроить, но установка его на значение 0,9 обычно работает хорошо. СК превосходит АдГрад и обычно работает лучше, чем алгоритм оптимизации импульса и УГН, поэтому он является одним из наиболее часто используемых оптимизаторов

$$S^{(n+1)} = \beta S^{(n)} + (1 - \beta) \left[\frac{\partial J}{\partial w} (w^{(n)}) \right]^2. \quad (13)$$

Оптимизация с адаптивной оценкой момента (АОМ) [40] сочетает в себе идеи АдГрад и СК: отслеживает как экспоненциально убывающее среднее значение прошлых градиентов, так и квадраты градиентов (как оптимизация импульса и СК соответственно), как показано в (14) и (15). Этот метод вводит два новых гиперпараметра для настройки, обычно β_1 устанавливают на 0,9, а β_2 на 0,999, что дает наилучший результат. Поскольку АОМ является адаптивным алгоритмом, ему требуется меньшая настройка скорости обучения η , и значение по умолчанию 0,001 также работает хорошо. Существование значений по умолчанию, которые доказали свою хорошую работу, делает оптимизацию АОМ более простой в использовании

$$\Delta^{(n+1)}_w = \frac{\eta}{\sqrt{S^{(n+1)} + \epsilon}} m^{(n+1)}, \quad (14)$$

где

$$m^{(n+1)} = \frac{\beta_1 m^{(n)} + (1 - \beta_1) \frac{\partial J}{\partial w}(w^{(n)})}{1 - \beta_1}; \quad (15)$$

$$S^{(n+1)} = \frac{\beta_2 S^{(n)} + (1 - \beta_2) \left[\frac{\partial J}{\partial w}(w^{(n)}) \right]^2}{1 - \beta_2}.$$

4.4.5. Ранняя остановка, регуляризация и отсев

Переобучение является серьезной проблемой для больших моделей и должно контролироваться. В этом разделе представлены основные методы предотвращения переобучения.

Чтобы оценить, насколько хорошо модель будет обобщаться, обычно используется часть набора данных (до 20%) из обучающих данных и оценивается предсказание модели на основе этого проверочного набора. Если гиперпараметры настраиваются с использованием этого проверочного набора, то следует зарезервировать дополнительный тестовый набор для проверки переобучения гиперпараметров.

Оценку ошибки обучения и проверки с учетом количества эпох можно использовать для отслеживания того, насколько хорошо модель работает в обоих наборах данных. Типичный график переобучения ИНС показан на рис. 22.

Когда начинается обучение, ошибки как обучения, так и проверки уменьшаются аналогичным образом. Однако по прошествии нескольких эпох ошибка обучения все еще уменьшается (хотя и медленно) по мере прохождения эпох, но ошибка проверки увеличивается, указывая на то, что модель перестраивается под данные в наборе данных.

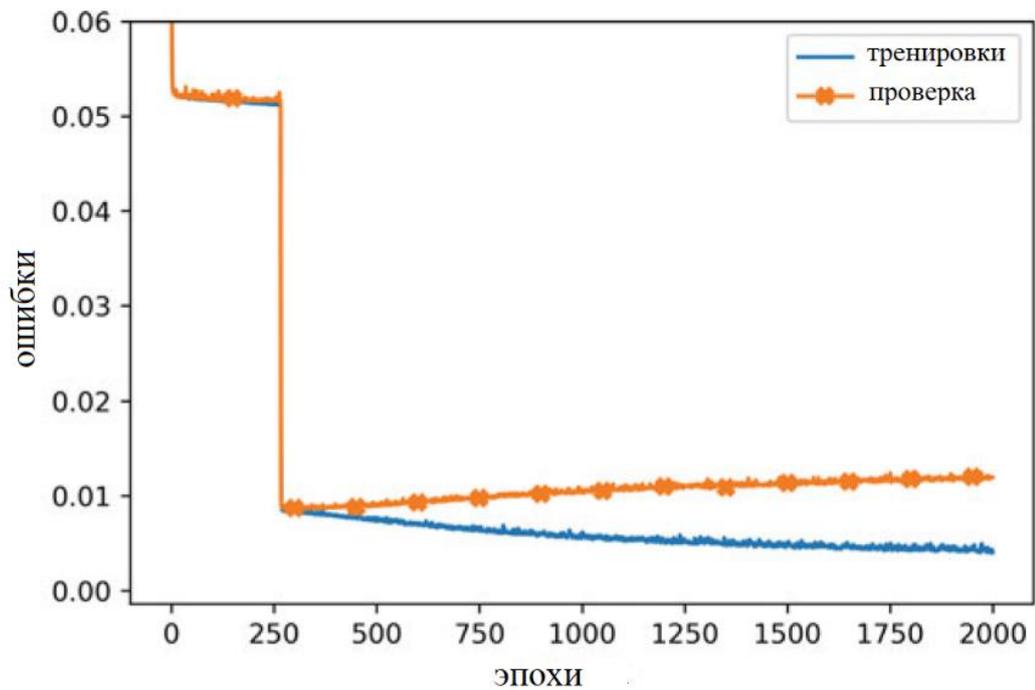


Рис. 22. Эволюция ошибок обучения и проверки

Ранняя остановка, как следует из названия, заключается в остановке обучения ИНС, как только в наборе проверки не будет улучшений. Чтобы реализовать раннюю остановку, систему можно моделировать по набору проверок с определенным интервалом эпох и сохранять моментальные снимки модели, которая имеет наилучшую производительность по сравнению с этим набором. Когда производительность при проверке продолжает ухудшаться в течение некоторых периодов, обучение может быть остановлено. Это простой подход для предотвращения переобучения; однако это также не позволяет проводить длительное обучение на сложных моделях.

Альтернативными методами регуляризации могут выступать методы лассо и гребня, которые ограничивают значение весов сети, увеличивая разреженность в модели, чтобы предотвратить переобучение. Реализованы регуляризация лассо и гребня через введение нового параметра в функцию потерь, используемую для обучения сети, как описано (16) и (17), соответственно, где $J'(w)$ – функция потерь, используемая для обучения сети, $J(w)$ – базовая функция потерь, w – вектор, содержащий все веса в ИНС, а λ – гиперпараметр, который управляет степенью регуляризации. $\|\cdot\|_1$ представляет норму L_1 , а $\|\cdot\|_2$ представляет норму L_2 . Из-за использования нормы L_1 и нормы L_2 методы регуляризации часто называют регуляризацией по L_1 и L_2 соответственно

$$J'(w|x, y) = J(w|x, y) + \lambda \sum |w_i| = J(w|x, y) + \lambda \|W\|_1 ; \quad (16)$$

$$J'(w|x, y) = J(w|x, y) + \lambda \sum w_i^2 = J(w|x, y) + \lambda \|W\|_2^2 . \quad (17)$$

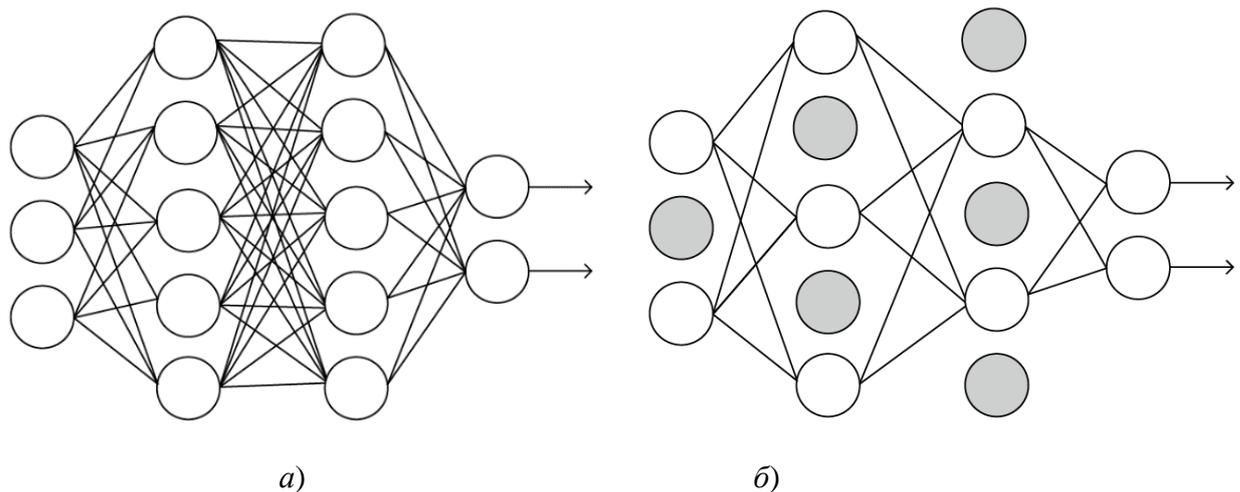


Рис. 23:

a – стандартная нейронная сеть; *б* – нейронная сеть после применения исключения, серые узлы представляют собой выпавшие единицы

Еще одним подходом для предотвращения переобучения является отсев [38]. Отсев состоит из отсева нейронов в нейронной сети путем временного удаления их из сети, как показано на рис. 23, *a*. Когда метод отсева применяется к ИНС, исходная выборка преобразуется, и в результате получается более простая сеть, как показано на рис. 23, *б*.

Эта выборка сети содержит только те нейроны, которые не были отсеяны, и поэтому ИНС с n нейронами соотносится с более простыми сетями 2^n (поскольку единицы нейронов могут быть активными или исключенными). Во время обучения для каждого шага обучения сеть отсеивается и обучается. В этом смысле применение отсева к сети эквивалентно обучению 2^n более простых сетей с разной архитектурой, где каждая из этих более простых сетей обучается очень редко, если вообще обучается. Объединение 2^n сетей с различными архитектурами в единую сеть, которая выдает среднее значение (аппроксимацию) выходных данных, приводит к значительно меньшей ошибке обобщения для широкого спектра задач [38].

Вероятность сохранения каждого узла p , является новым гиперпараметром, который определяет, насколько сильно применяется отсев. Обычно используется 50 %-ная вероятность того, что узел будет активен. Прежде чем использовать сеть для составления прогнозов, исходящие веса каждой единицы умножаются на вероятность p , используемую во время обучения.

4.5. ВЫВОД

Таким образом были рассмотрены основные составляющие машинного обучения. ИНС – это очень гибкие модели, которые могут быть легко адаптированы к целевой задаче и обладают способностью смешивать задачи

регрессии и классификации в одной модели. Кроме того, они способны обрабатывать данные с высокой пропускной способности. Хотя время их вычислений может варьироваться в зависимости от объема обрабатываемых данных и быть довольно длительным, это не влияет на способность к прогнозированию, поскольку наиболее трудоемкой задачей является обучение. Гибкость ИНС заключается в том, что существует множество требующих настройки гиперпараметров, наилучшие значения которых зависят от проблемы и подручных средств и могут быть найдены только методом проб и ошибок. В главе были представлены наиболее релевантные гиперпараметры, а также некоторые стратегии выбора и обучения наилучшей модели.

5. ИНФОРМАЦИОННЫЕ СИСТЕМЫ И РЕСУРСЫ СЕТИ ИНТЕРНЕТ ДЛЯ ПОИСКА ПРАВОВОЙ ИНФОРМАЦИИ

В современном мире своевременный доступ к достоверной актуальной информации играет решающее значение во многих сферах человеческой деятельности. Известное выражение гласит: «Кто владеет информацией, тот владеет миром». Информация влияет на судьбы людей. Так, всего лишь отсутствие запятой приводит к неопределенности: «Казнить нельзя помиловать». А возможная ошибка в постановке запятой приведет к летальным последствиям: «Казнить, нельзя помиловать». Ко всему прочему «незнание законов не освобождает от ответственности».

Использование современных ИТ и возможностей (в том числе и сетевых) ИС позволяет обеспечить эффективный доступ к правовой информации, как профессиональных юристов, так и специалистов других специальностей, которые должны учитывать различные правовые аспекты в своей деятельности, а также студентов, которые в процессе своего обучения закладывают основы информационной культуры при работе с такими информационными технологиями и системами. И, конечно, всем нам «обычным обывателям» важно «не потеряться» в огромных массивах постоянно меняющейся правовой информации.

Важно, что Конституция РФ в ст. 15 закрепляет обязательность официальной публикации законов (см. рис. 24).

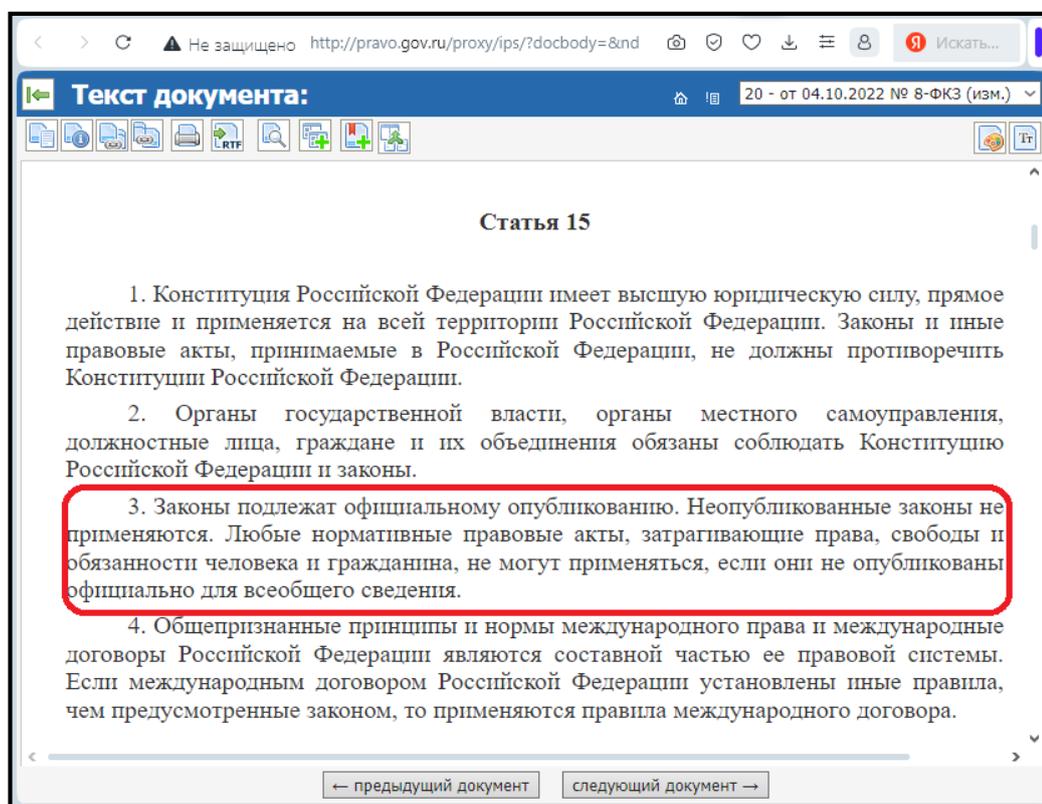


Рис. 24. Конституция РФ (ст. 15) на странице официального интернет-портала правовой информации

Следует отметить, что при использовании разного рода как коммерческих, так и некоммерческих информационных систем, предоставляющих правовую информацию необходимо учитывать, что круг официальных источников опубликования весьма ограничен (например, определяется Указом Президента РФ от 23 мая 1996 г. № 763 «О порядке опубликования и вступления в силу актов Президента Российской Федерации, Правительства Российской Федерации и нормативных правовых актов федеральных органов исполнительной власти»).

Тексты правовых актов с внесенными в них изменениями размещаются на «Официальном интернет-портале правовой информации» (www.pravo.gov.ru) в соответствии с Указом Президента Российской Федерации от 3 марта 2022 № 90 «О некоторых вопросах размещения текстов правовых актов на «Официальном интернет-портале правовой информации. (www.pravo.gov.ru)»).

Официальными являются также тексты актов Президента Российской Федерации, в том числе с внесенными в них изменениями, и актов Правительства Российской Федерации, распространяемые в электронном виде федеральным государственным унитарным предприятием «Научно-технический центр правовой информации «Система Федеральной службы охраны Российской Федерации», а также органами государственной охраны.

Также официальным является опубликование в «Российской газете», Собрании законодательства Российской Федерации и на «Официальном интернет-портале правовой информации», функционирование которого обеспечивает Федеральная служба охраны Российской Федерации.

Главная страница официального интернет-портала правовой информации www.pravo.gov.ru (см. рис. 25) дает различные возможности поиска правовой информации: «Официальное опубликование» (см. рис. 26), «Тексты правовых актов с внесенными изменениями» (см. рис. 27), «Законодательство России», «Новые поступления» и др.

Таким образом, наличие официальных источников публикации нормативных и правовых актов позволяет при необходимости свободно (бесплатно) получить доступ к актуальной и достоверной правовой информации, проверить достоверность информации, полученной из других источников.

Почему же, несмотря на наличие таких официальных источников правовой информации, все большей популярностью пользуются коммерческие правовые информационные системы, например такие как «Гарант» (www.garant.ru), «КонсультантПлюс» (www.consultant.ru) и ряд других? Все дело в том, что такие системы обладают большими функциональными возможностями, нежели просто поиск нормативных правовых актов. Это весьма эффективный инструмент как для профессионала, так и для обычного пользователя. Далее мы рассмотрим некоторые особенности и основы работы с наиболее известными и распространенными в нашей стране правовыми информационными системами «Гарант» и «Консультант Плюс». Эти системы более чем за 30 лет своего существования по праву завоевали большую часть рынка. Каждая из них шла своим путем, ставила свои цели, но несмотря

на некоторые отличия они имеют весьма продуманный удобный пользовательский интерфейс, большой объем накопленной нормативной правовой и иной информации, а также в чем-то сходную функциональную структуру [28].

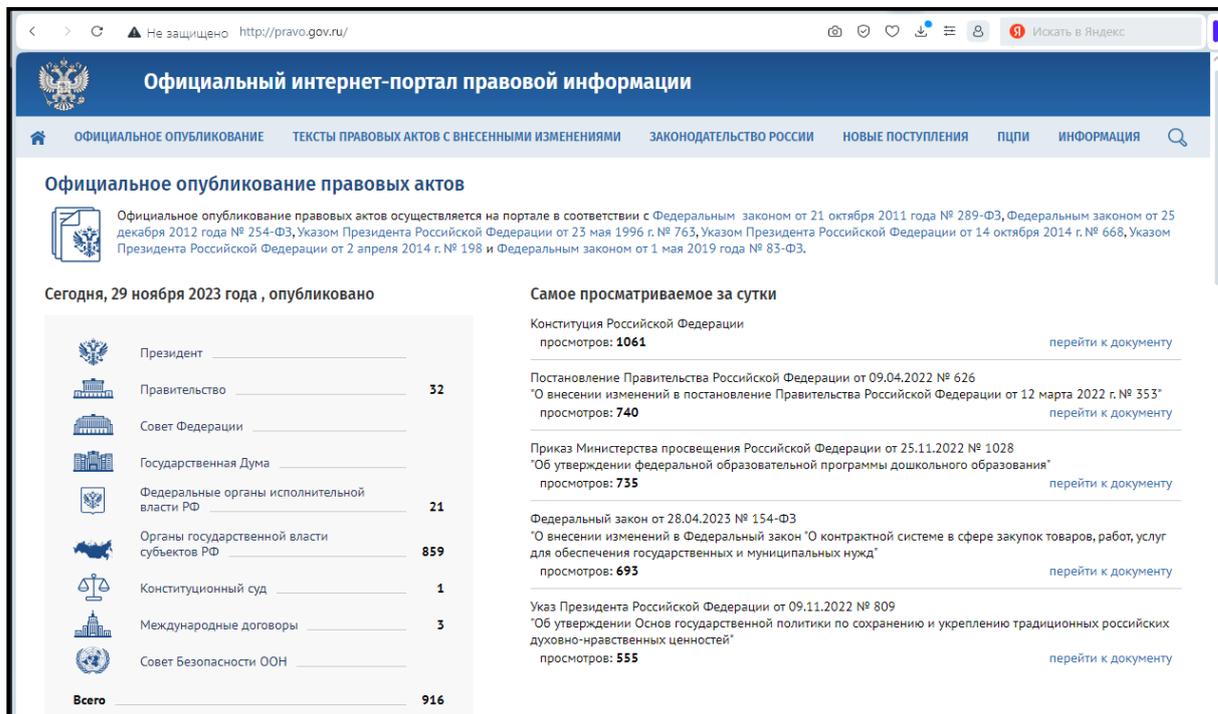


Рис. 25. Главная страница официального интернет-портала правовой информации

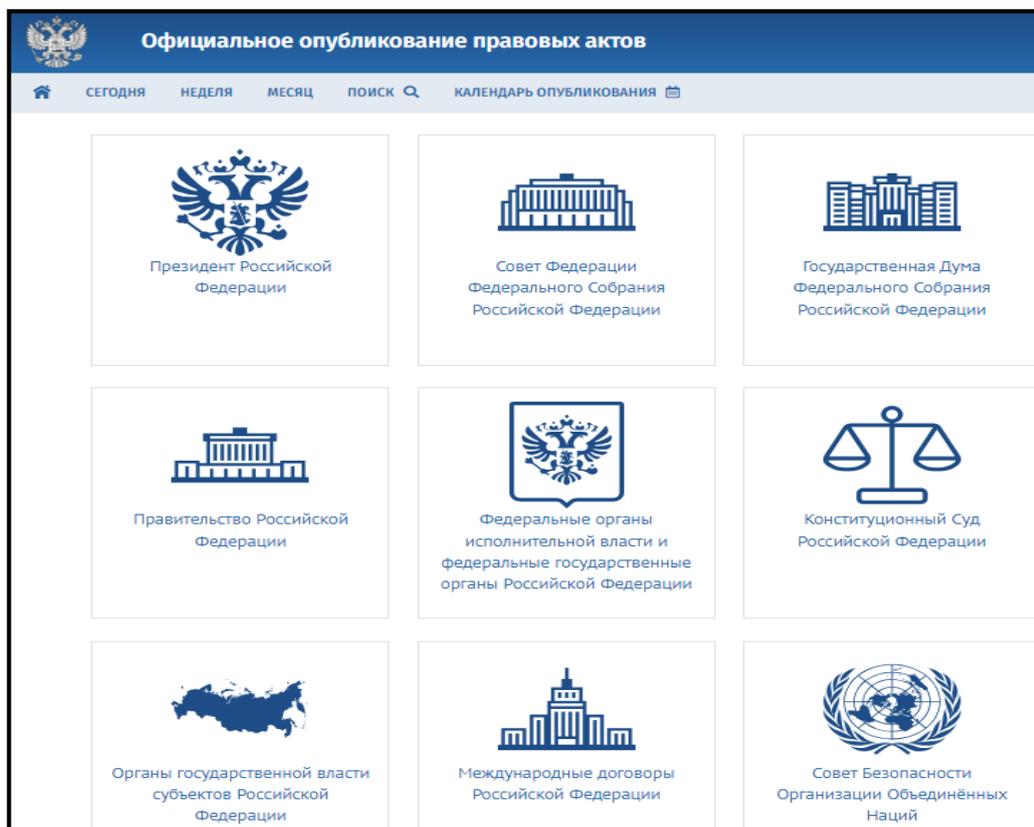


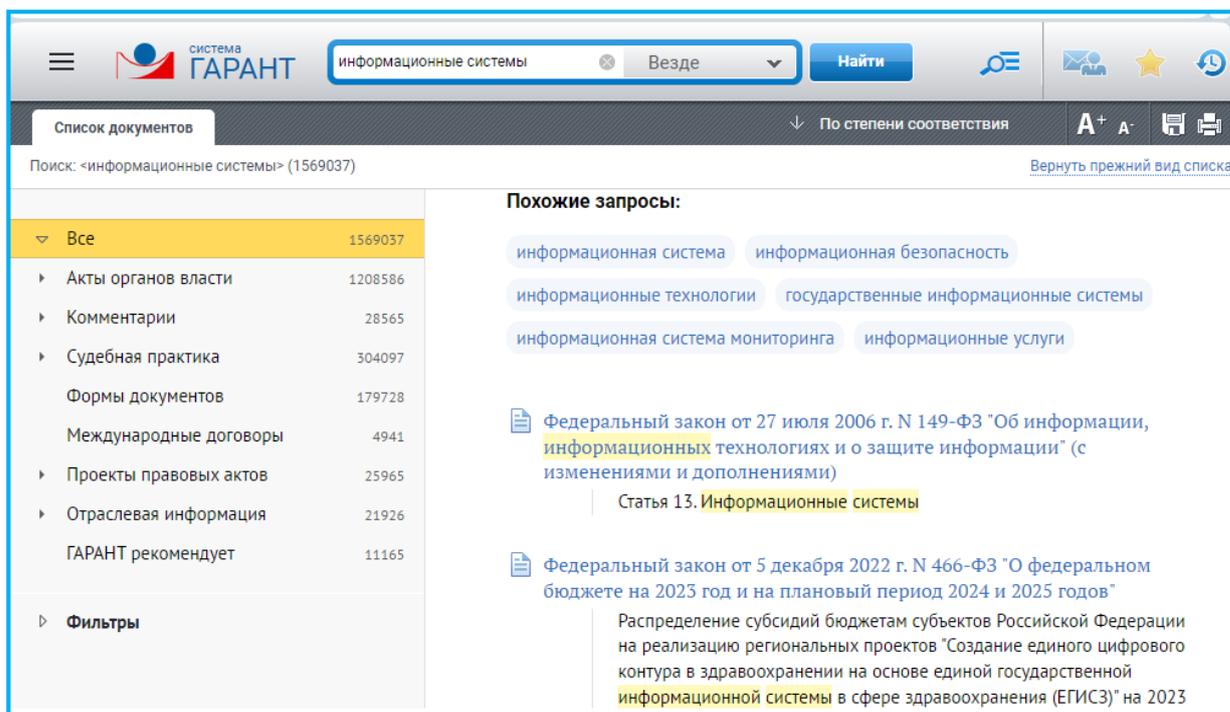
Рис. 26. Страница «Официальное опубликование правовых актов» официального интернет-портала правовой информации

Рис. 27. Страница «Официальное опубликование» официального интернет-портала правовой информации (карточка запроса по реквизитам документов)

Типичная система правовой информации включает в себя:

- средства поиска документов по контексту и рубрикатору;
- средства поиска документа по реквизитам;
- механизм навигации в базе данных по гипертекстовым ссылкам;
- модули работы со списками и текстами документов;
- подсистему обновления базы данных.

Информационная система «Гарант» имеет разнообразные возможности поиска требуемой информации. Вариант выбранного инструмента поиска определяется в первую очередь тем, насколько полной информацией о требуемом документе вы владеете. Шутка «ну ты и ответил... ну ты и спросил» как нельзя полно характеризует результат поискового запроса. Чем большей предварительной информацией мы владеем, чем точнее сформулирован запрос, тем больше будет удовлетворенность результатами запроса. Хотя, следует отметить, что и при наличии весьма ограниченной информации система обычно дает приемлемый результат. Это видно на примере использования строки «Базовый поиск», где был набран запрос «информационные системы» и поучен список документов, включая Федеральный закон № 149-ФЗ, в котором в ст. 13 дается искомое определение. Если требуется уточнить запрос, то можно воспользоваться всевозможными доступными фильтрами («Все», «Акты органов власти» и т.д.). И если первые версии системы требовали набора текста в именительном падеже, то современные – ищут информацию при произвольном варианте набора.



**Рис. 28. Интерфейс системы «ГАРАНТ»
(на примере использования строки «Базовый поиск»)**

Интерфейс системы «ГАРАНТ» (на примере использования строки «Базовый поиск»), приведен ниже (см. рис. 28).

В системе имеются различные виды поиска:

- базовый поиск;
- поиск по реквизитам;
- поиск с указанием точных реквизитов;
- поиск по судебной практике;
- дополнительные виды поиска: расширенный поиск (поиск с описанием реальной ситуации), поиск по публикации, поиск лекарственных средств.

Нажав на элемент , который находится правее строки базового поиска, можно получить доступ к расширенному поиску (по реквизитам, по судебной практике, по ситуации и т.д.) [13].

Поиск по реквизитам – это точный инструмент, предназначенный для поиска документов по заранее известной (или предполагаемой) информации о документе. Он позволяет сочетать в запросе самую разнообразную информацию: тип и номер искомого документа, принявший орган и дату принятия, опубликования или регистрации в Минюсте, слова или словосочетания, содержащиеся в тексте документа, и многие другие реквизиты.

Каждый документ характеризуется основными и расширенными реквизитами, значения которых задаются в качестве условий поиска. Для результативного поиска достаточно заполнить от одного до трех полей карточки запроса. Вариант запроса при поиске по реквизитам, например, ищем уголовный кодекс, установив фильтр «действующий».

Поиск по ситуации – уникальная запатентованная разработка компании «Гарант». Энциклопедия ситуаций «Гарант» содержит свыше 100 000 подробных терминов («Гарант-Максимум. Вся Россия»). Благодаря Энциклопедии ситуаций любой человек, не являющийся специалистом в рассматриваемой области и не знающий реквизитов нормативного акта, может отыскать в огромном массиве данных ответ на свой вопрос, подобрать документы, описывающие ситуацию. Используя привычные термины, формулируют вопросы, и в поле контекстного фильтра вводят в любой последовательности начальные части слов. В результате, все ситуации, названия которых соответствуют набранному контексту, будут наглядно представлены в основном окне системы.

Вариант запроса при поиске по ситуации, например, ищем «налог транспортный на физических лиц». После выполнения команды будут построены только действующие и наиболее важные документы.

Поиск по публикации производится, если требуется найти авторские материалы.

Вариант запроса при поиске по источнику опубликования статьи, например, ищем «Споры о государственной регистрации прав на недвижимость» (А. Яковлева, «Законодательство», № 1, январь 2005 г.).

Поиск по словарю нормативных определений выполняется во встроенном в «Гарант» толковом словаре терминов. Словарь содержит толкования десятков тысяч терминов экономической и правовой тематики. Источниками являются отечественные и зарубежные энциклопедии, специализированные справочники, толковые словари и определения из нормативных актов.

Вариант запроса при поиске по словарю понятия, например, «компенсация морального вреда».

В информационной системе «Гарант» имеется механизм навигации в базе данных по гипертекстовым ссылкам.

На Главной странице «Гарант», помимо вкладки Главная (где доступны все кодексы РФ, справочная информация и др.), также доступны вкладки Новости ПРАЙМ, Энциклопедия решений, Все решения Гаранта (где представлены аналитика, специальные базы данных, услуги). Раздел Справочная информация основного меню используется для быстрого поиска справочной информации (формы отчетности, ставки налогов, курсы валют, адреса и телефоны органов власти и многое другое) [13].

Информационная система «Гарант» позволяет выполнять следующие операции (наиболее удобно, по нашему мнению, их выполнять с помощью контекстного меню для выбранного объекта по нажатию правой клавиши мыши).

Работа со списками документов:

- просмотр списка документов (список которых появляется, например, после выполнения команды);
- поиск контекста в названиях документов списка (Ctrl + F);

- дополнительный поиск в активном окне (через карточку поиска по реквизитам);
- синхронный просмотр;
- логические операции со списком (сортировка, фильтрация, редактирование и т.д.);
- сохранение в папки;
- экспорт и печать (статьи, книги, энциклопедии) и, если известно, где была опубликована интересующая статья.

Работа с текстами документов:

- структура документа (оглавление, элементы текста, цветовое выделение, гиперссылки, комментарии);
- все доступные операции имеются в контекстном меню;
- анализ документа (F8 – справка к документу; Alt F8 – корреспонденты к документу; Ctrl F8 – респонденты к документу);
- графическая копия официальной публикации;
- работа с извлечением. Режим просмотра текста документа «в извлечениях» позволяет скрывать те его разделы, которые не имеют отношения к тематике предшествующего поиска;
- «машина времени» – позволяет посмотреть редакцию документа на заданную дату.

Операции с документами (все доступные операции имеются в контекстном меню):

- печать документа;
- сохранение в файл;
- экспорт в MS Word, Excel;
- сохранение в папку, установка закладки;
- постанова на контроль (позволяет оперативно получать уведомления об изменении документа на контроле).

Необходимо отметить, что в системе «Гарант» работает «Журнал работ», протоколирующий все запросы.

На сайте www.garant.ru можно найти всю необходимую информацию о «Гарант», а также весьма подробно изучить все приемы работы с этой системой на практике. Можно также пройти дистанционное тестирование на знание системы и использования ее возможностей [13].

На Главной странице «ГАРАНТ» доступна вкладка «Все решения Гаранта» (см. рис. 29).

На вкладке Продукты и услуги/Информационные материалы (<https://www.garant.ru/products/flipbook/>), расположенной на сайте «Гарант», имеются необходимые для изучения материалы (практическое руководство, решение практических ситуаций, каталог продуктов и услуг и др.).

На сайте <http://learning.garant.ru/> доступны обучающие видео по работе с системой.

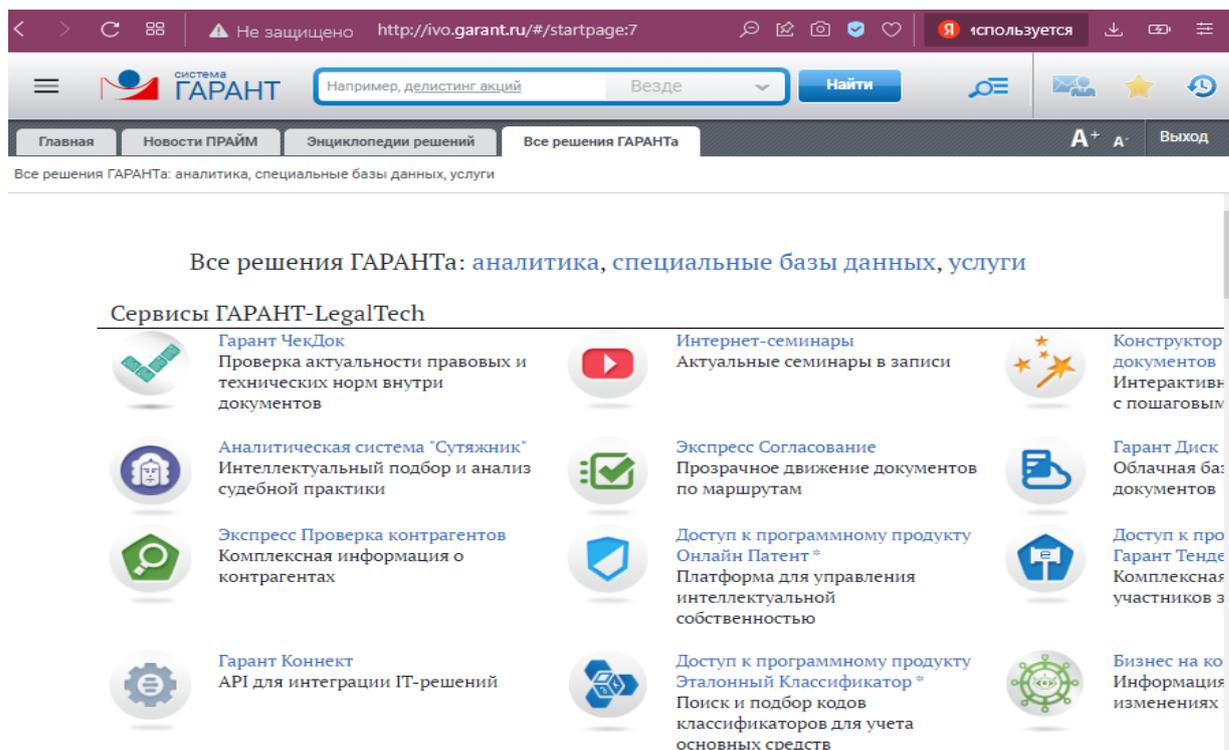


Рис. 29. Главная страница «ГАРАНТ» (вкладка все решения Гаранта)

Не менее важным ресурсом для поиска правовой информации в сети Интернет является официальный сайт компании «КонсультантПлюс»: (www.consultant.ru) (см. рис. 30).

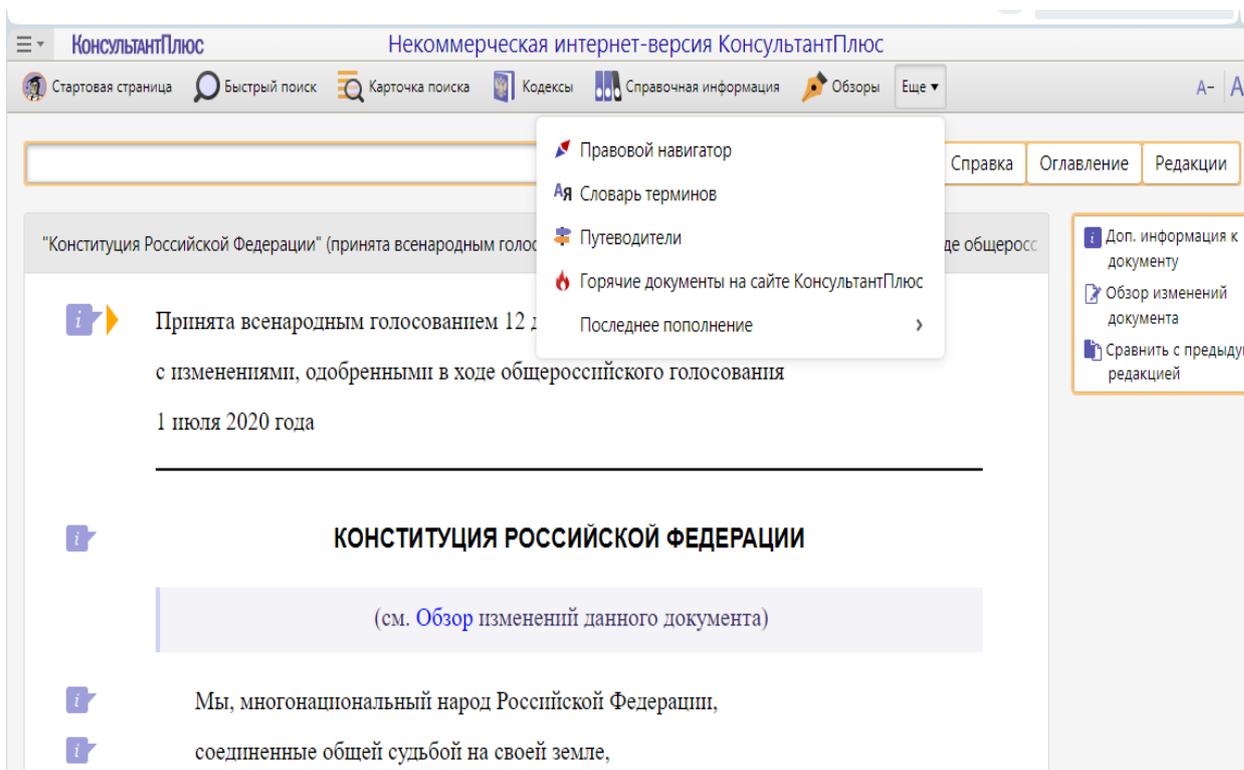


Рис. 30. Фрагмент страницы для поиска правовой информации с официального сайта компании «КонсультантПлюс»

Система «КонсультантПлюс» также как и «Гарант» уже более 30 лет предлагает свои решения (в 2022 г. сеть КонсультантПлюс отметила 30-летие) [18].

Стоит отметить, что базовые возможности поиска правовой информации похожи на те, которые мы рассмотрели в «Гарант».

Например, в «Гарант» мы рассматривали строку «Базовый поиск», то в «КонсультантПлюс» – это строка «Быстрого поиска». «Карточка поиска» в «КонсультантПлюс» чем-то напоминает «Поиск по реквизитам» в «Гарант», а «Правовой навигатор» в «КонсультантПлюс» отчасти похож на «Поиск по ситуации» в «Гарант» и т.д.

Конечно же, есть у каждого продукта и свои особенности. Более подробно с ними и с тонкостями работы можно ознакомиться на сайте <https://www.consultant.ru/about/support/> на странице «В помощь пользователю». Там же можно скачать руководство пользователя «Шаг за шагом» https://static.consultant.ru/obj/file/study/cons_manual_2023.rar (см. рис. 31).

Стоит отметить что и «Гарант» и «КонсультантПлюс» (с несколько урезанным функционалом) доступны к использованию на соответствующих сайтах и в некоммерческих версиях. Регистрация пользователя дает возможность ознакомиться с более полным функционалом в течение нескольких дней.

Еще важно помнить, что в случае недоступности правовых документов или в случае сомнений по поводу их актуальности или достоверности при работе с нормативно-правовыми документами в сети Интернет всегда доступен Официальный интернет-портал правовой информации (<http://pravo.gov.ru/>), представляющий собой орган официального опубликования в Российской Федерации федеральных конституционных законов, федеральных законов, актов палат Федерального собрания, указов и распоряжений Президента Российской Федерации.

В помощь пользователю

Это раздел предназначен для всех, кто хочет совершенствовать свои навыки работы с системой КонсультантПлюс.

Новички найдут здесь материалы, которые помогут освоить КонсультантПлюс "с нуля". Опытные пользователи смогут значительно расширить полезные навыки поиска ответов на профессиональные вопросы с помощью КонсультантПлюс.

Рекомендуем тем, кто только знакомится с системой КонсультантПлюс

Руководство пользователя "КонсультантПлюс: Шаг за шагом"



Руководство пользователя поможет научиться эффективно использовать все возможности системы КонсультантПлюс. Наглядные примеры, пошаговые инструкции, проработка ситуаций, с которыми вы можете столкнуться в вашей работе, - все это позволит быстро освоить приемы работы с КонсультантПлюс.

Материал предназначен как для начинающих, так и для имеющих опыт работы с системой пользователей.

Скачать руководство пользователя "Шаг за шагом" (97 Мб)

Рис. 31. Фрагмент страницы «В помощь пользователю» с официального сайта компании «КонсультантПлюс»

6. ПРАВОВЫЕ ОСНОВЫ ЗАЩИТЫ ПЕРСОНАЛЬНЫХ ДАННЫХ ПРИ ИСПОЛЬЗОВАНИИ СОВРЕМЕННЫХ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ И СИСТЕМ

6.1. НОРМАТИВНЫЕ ПРАВОВЫЕ АКТЫ В ОБЛАСТИ ПЕРСОНАЛЬНЫХ ДАННЫХ

Под персональными данными обычно подразумевают различные сведения о физическом лице, такие как его фамилия, имя, отчество, год, месяц, дата и место рождения, адрес, семейное, социальное, имущественное положение, образование, профессия, доходы и другую информацию.

Одним из принципов, лежащих в основе всех действий, связанных с оборотом информации, должен быть принцип законности. Поэтому важным является наличие соответствующей правовой основы в рассматриваемой сфере.

Правовую основу обращения с персональными данными физических лиц в целях реализации конституционных прав человека, в том числе права на неприкосновенность частной жизни, личную и семейную тайну составляет законодательство Российской Федерации в области персональных данных, которое основывается на Конституции Российской Федерации и международных договорах Российской Федерации и состоит из Федерального закона от 27 июля 2006 г. № 152-ФЗ «О персональных данных» и других определяющих случаи и особенности обработки персональных данных федеральных законов.

В Федеральном законе от 27.07.2006 г. № 152-ФЗ «О персональных данных» определены основные понятия, связанные с персональными данными. Приведем ниже некоторые из них [24].

Персональные данные – любая информация, относящаяся прямо или косвенно к определяемому физическому лицу (субъекту персональных данных).

Персональные данные, разрешенные субъектом персональных данных для распространения, – персональные данные, доступ неограниченного круга лиц к которым предоставлен субъектом персональных данных путем дачи согласия на обработку персональных данных, разрешенных субъектом персональных данных для распространения в порядке, предусмотренном Федеральным законом.

Оператор – государственный орган, муниципальный орган, юридическое или физическое лицо, самостоятельно или совместно с другими лицами организующее и(или) осуществляющее обработку персональных данных, а также определяющее цели обработки персональных данных, состав персональных данных, подлежащих обработке, действия (операции), совершаемые с персональными данными.

Обработка персональных данных – любое действие (операция) или совокупность действий (операций), совершаемых с использованием средств автоматизации или без использования таких средств с персональными данными, включая сбор, запись, систематизацию, накопление, хранение, уточнение (обновление, изменение), извлечение, использование, передачу (распространение, предоставление, доступ), обезличивание, блокирование, удаление, уничтожение персональных данных.

Автоматизированная обработка персональных данных – обработка персональных данных с помощью средств вычислительной техники.

Информационная система персональных данных – совокупность содержащихся в базах данных персональных данных и обеспечивающих их обработку информационных технологий и технических средств.

Распространение персональных данных – действия, направленные на раскрытие персональных данных неопределенному кругу лиц.

Предоставление персональных данных – действия, направленные на раскрытие персональных данных определенному лицу или определенному кругу лиц.

Блокирование персональных данных – временное прекращение обработки персональных данных (за исключением случаев, если обработка необходима для уточнения персональных данных).

Уничтожение персональных данных – действия, в результате которых становится невозможным восстановить содержание персональных данных в информационной системе персональных данных и (или) в результате которых уничтожаются материальные носители персональных данных.

Обезличивание персональных данных – действия, в результате которых становится невозможным без использования дополнительной информации определить принадлежность персональных данных конкретному субъекту персональных данных.

Трансграничная передача персональных данных – передача персональных данных на территорию иностранного государства органу власти иностранного государства, иностранному физическому лицу или иностранному юридическому лицу.

Знание определений используемых терминов позволяет правильно понимать и исполнять требования нормативных и правовых актов в этой сфере.

Весьма важно отслеживать и изменения в нормативных актах, в чем существенную помощь оказывают информационные ресурсы и информационные сообщения на сайтах таких ведомств, как Роскомнадзор (rkn.gov.ru), ФСТЭК – Федеральная служба по техническому и экспортному контролю (fstec.ru), ФСБ – Федеральная служба безопасности Российской Федерации, pravo.gov.ru – Официальный интернет-портал правовой информации и др.

Следует отметить, что к настоящему времени наработана весьма обширная правовая база, регулирующая работу с персональными данными. Отвечая

на запросы времени, она постоянно уточняется и дополняется. Некоторые из таких документов, включая и Федеральный закон от 27.07.2006 г. № 152-ФЗ «О персональных данных», приведены ниже [24].

Нормативные правовые акты в области персональных данных.

- Федеральный закон от 27.07.2006 г. № 152-ФЗ «О персональных данных»;
- Постановление Правительства Российской Федерации от 01.11.2012 г. № 1119 «Об утверждении требований к защите персональных данных при их обработке в информационных системах персональных данных»;
- Приказ Федеральной службы по надзору в сфере связи, информационных технологий и массовых коммуникаций от 28.10.2022 № 180 «Об утверждении форм уведомлений о намерении осуществлять обработку персональных данных, об изменении сведений, содержащихся в уведомлении о намерении осуществлять обработку персональных данных, о прекращении обработки персональных данных»;
- Приказ Федеральной службы по надзору в сфере связи, информационных технологий и массовых коммуникаций от 28.10.2022 № 179 «Об утверждении Требований к подтверждению уничтожения персональных данных»;
- Приказ Федеральной службы по надзору в сфере связи, информационных технологий и массовых коммуникаций от 27.10.2022 № 178 «Об утверждении Требований к оценке вреда, который может быть причинен субъектам персональных данных в случае нарушения Федерального закона «О персональных данных»;
- Постановление Правительства Российской Федерации от 29.06.2021 № 1046 «О федеральном государственном контроле (надзоре) за обработкой персональных данных»;
- Постановление Правительства РФ от 21.03.2012 № 211 (ред. от 15.04.2019) «Об утверждении перечня мер, направленных на обеспечение выполнения обязанностей, предусмотренных Федеральным законом «О персональных данных» и принятыми в соответствии с ним нормативными правовыми актами, операторами, являющимися государственными или муниципальными органами»;
- Конвенция о защите физических лиц при автоматизированной обработке персональных данных (Страсбург, 28 января 1981 г.);
- Директива Европейского Союза № 2002/58/ЕС «О приватности и электронных коммуникациях»;
- Трудовой кодекс Российской Федерации от 30.12.2001 г. № 197-ФЗ – Глава 14 «Защита персональных данных работника»;
- Федеральный закон от 19.12.2005 г. № 160-ФЗ «О ратификации Конвенции Совета Европы о защите физических лиц при автоматизированной обработке персональных данных»;

- Федеральный закон от 27.07.2006 г. № 149-ФЗ «Об информации, информационных технологиях и о защите информации»;
- Федеральный закон Российской Федерации от 25.07.2011 г. № 261-ФЗ «О внесении изменений в Федеральный закон «О персональных данных»;
- Федеральный закон от 30.12.2015 г. № 439-ФЗ «О внесении изменений в Кодекс Российской Федерации об административных правонарушениях»;
- Федеральный закон от 21.07.2014 г. № 242-ФЗ «О внесении изменений в отдельные законодательные акты Российской Федерации в части уточнения порядка обработки персональных данных в информационно-телекоммуникационных сетях»;
- Указ Президента Российской Федерации от 06.03.1997 г. № 188 «Об утверждении перечня сведений конфиденциального характера»;
- Указ Президента Российской Федерации от 30.05.2005 г. № 609 «Об утверждении Положения о персональных данных государственного гражданского служащего Российской Федерации и ведении его личного дела»;
- Указ Президента Российской Федерации от 17.03.2008 г. № 351 «О мерах по обеспечению информационной безопасности Российской Федерации при использовании информационно-телекоммуникационных сетей международного информационного обмена»;
- Распоряжение Президента Российской Федерации от 10.07.2001 г. № 366-РП «О подписании Конвенции о защите физических лиц при автоматизированной обработке персональных данных»;
- Постановление Правительства Российской Федерации от 21.03.2012 г. № 211 «Об утверждении перечня мер, направленных на обеспечение выполнения обязанностей, предусмотренных Федеральным законом «О персональных данных» и принятыми в соответствии с ним нормативными правовыми актами, операторами, являющимися государственными или муниципальными органами»;
- Постановление Правительства Российской Федерации от 03.11.1994 г. № 1233 «Об утверждении положения о порядке обращения со служебной информацией ограниченного распространения в федеральных органах исполнительной власти, уполномоченном органе управления использования атомной энергии и уполномоченном органе по космической деятельности»;
- Постановление Правительства Российской Федерации от 06.07.2008 г. № 512 «Об утверждении требований к материальным носителям биометрических персональных данных и технологиям хранения таких данных вне информационных систем персональных данных»;
- Постановление Правительства Российской Федерации от 15.09.2008 г. № 687 «Об утверждении положения об особенностях обработки персональных данных, осуществляемой без использования средств автоматизации»;

- Постановление Правительства РФ от 04.03.2010 г. № 125 «О перечне персональных данных, записываемых на электронные носители информации, содержащиеся в основных документах, удостоверяющих личность гражданина Российской Федерации, по которым граждане Российской Федерации осуществляют выезд из Российской Федерации и въезд в Российскую Федерацию»;
- Приказ Роскомнадзора от 05.09.2013 г. № 996 «Об утверждении требований и методов по обезличиванию персональных данных»;
- Распоряжение Правительства Российской Федерации от 15.08.2007 г. № 1055-Р «О плане подготовки проектов нормативных актов, необходимых для реализации Федерального закона «О персональных данных»;
- Приказ ФСБ России от 09.02.2005 г. № 66 «Об утверждении Положения о разработке, производстве, реализации и эксплуатации шифровальных (криптографических) средств защиты информации. Положение ПКЗ 2005»;
- Приказ ФСТЭК России от 18.02.2013 г. № 21 «Об утверждении Составы и содержания организационных и технических мер по обеспечению безопасности персональных данных при их обработке в информационных системах персональных данных»;
- Приказ Минкомвязи России от 20.07.2017 г. № 373 «О признании утратившими силу приказов Министерства связи и массовых коммуникаций РФ» от 21 декабря 2011 № 346, от 28 августа 2015 № 315 и п. 9 приказа Министерства связи и массовых коммуникаций РФ от 24 ноября 2014 № 403.

Важно отметить, что в соответствии с ч. 3 ст. 17 Конституции РФ «любые нормативные правовые акты, затрагивающие права, свободы и обязанности человека и гражданина, не могут применяться, если они не опубликованы официально для всеобщего сведения [17].

В соответствии с Конституцией России (ч. 2 ст. 55) в Российской Федерации не должны издаваться законы, отменяющие или умаляющие права и свободы человека и гражданина: «Права и свободы человека и гражданина могут быть ограничены федеральным законом только в той мере, в какой это необходимо в целях защиты основ конституционного строя, нравственности, здоровья, прав и законных интересов других лиц, обеспечения обороны страны и безопасности государства» [17].

6.2. НЕКОТОРЫЕ ОСОБЕННОСТИ ОБРАБОТКИ ПЕРСОНАЛЬНЫХ ДАННЫХ

Федеральный закон от 27.07.2006 г. № 152-ФЗ «О персональных данных», определяя принципы и условия обработки персональных данных, устанавливает общий запрет на обработку персональных данных без согласия субъекта персональных данных, закон предусматривает случаи, когда такое

согласие не требуется. Отдельно регулируются отношения по обработке специальных категорий персональных данных (сведения о расовой, национальной принадлежности, политических взглядах, религиозных или философских убеждениях, состоянии здоровья, интимной жизни). Обработка указанных категорий сведений не допускается без предварительного согласия субъекта персональных данных, за исключением случаев, когда персональные данные являются общедоступными, обработка данных необходима для обеспечения жизни и здоровья лица; обработка проводится в связи с осуществлением правосудия, а также иных обстоятельств [28].

Принципы обработки персональных данных.

1. Обработка персональных данных должна осуществляться на законной и справедливой основе.

2. Обработка персональных данных должна ограничиваться достижением конкретных, заранее определенных и законных целей. Не допускается обработка персональных данных, несовместимая с целями сбора персональных данных.

3. Не допускается объединение баз данных, содержащих персональные данные, обработка которых осуществляется в целях, несовместимых между собой.

4. Обработке подлежат только персональные данные, которые отвечают целям их обработки.

5. Содержание и объем обрабатываемых персональных данных должны соответствовать заявленным целям обработки. Обрабатываемые персональные данные не должны быть избыточными по отношению к заявленным целям их обработки.

6. При обработке персональных данных должны быть обеспечены точность персональных данных, их достаточность, а в необходимых случаях и актуальность по отношению к целям обработки персональных данных. Оператор должен принимать необходимые меры либо обеспечивать их принятие по удалению или уточнению неполных или неточных данных.

7. Хранение персональных данных должно осуществляться в форме, позволяющей определить субъекта персональных данных, не дольше, чем этого требуют цели обработки персональных данных, если срок хранения персональных данных не установлен федеральным законом, договором, стороной которого, выгодоприобретателем или поручителем, по которому является субъект персональных данных. Обрабатываемые персональные данные подлежат уничтожению либо обезличиванию по достижении целей обработки или в случае утраты необходимости в достижении этих целей, если иное не предусмотрено федеральным законом.

Важнейшей гарантией прав субъекта персональных данных является обязанность операторов и третьих лиц, получивших доступ к персональным

данным, обеспечивать их конфиденциальность (кроме случаев их обезличивания и общедоступных персональных данных), а также право субъекта персональных данных на защиту своих прав и законных интересов, в том числе на возмещение убытков и(или) компенсацию морального вреда в судебном порядке.

Контроль и надзор за обработкой персональных данных возложен на федеральный орган исполнительной власти, осуществляющий функции по контролю и надзору в сфере информационных технологий и связи, который наделяется соответствующими правами и обязанностями. В частности, уполномоченный орган вправе осуществлять проверку информационной системы обработки персональных данных, предъявлять требования по блокированию, удалению недостоверных или полученных незаконным путем персональных данных, устанавливать постоянный или временный запрет на обработку персональных данных, проводить расследования в порядке административного производства о нарушениях закона.

Федеральный закон № 152-ФЗ устанавливает принципы трансграничной передачи данных, при которой должна обеспечиваться адекватная защита прав субъектов персональных данных.

Следует отметить, что в 2022 году в законе «О персональных данных» появились положения, которые ограничивают сбор персональных данных, если они не являются необходимыми для исполнения договора. Введена добровольная сдача биометрических данных. Такие же ограничения установлены и законом о защите прав потребителей, введена административная ответственность за нарушение этих норм [24].

Недопустимо, когда операторы персональных данных запрашивают избыточные сведения. Потребители часто сталкиваются с тем, что личные сведения становятся разменной монетой. Интерес к персональным данным потребителя нередко превалирует над интересом заключить договор.

Еще одно нововведение в законодательстве – это запрет на использование иностранных мессенджеров для передачи платежных документов или предоставления информации, которая содержит персональные данные. Их перечень утвержден Роскомнадзором и размещен на сайте ведомства.

Вот лишь некоторые важные изменения, которые необходимо учитывать при обработке персональных данных (ПД) с 1 марта 2023 года.

Новый порядок направления уведомлений об изменениях.

Согласно новым правилам, оператор ПД обязан уведомить Роскомнадзор об изменениях, представленных им ранее сведений об обработке ПД, произошедших за месяц, в срок не позднее 15 числа следующего месяца (ч. 7 ст. 22 Федерального закона от 27 июля 2006 г. № 152-ФЗ «О персональных данных»). Подается уведомление в отношении всех изменений.

За невыполнение требований организацию могут привлечь к административной ответственности по ст. 19.7 КоАП РФ – оштрафовать на сумму до 5 тыс. р. [16].

Обязанность хранить доказательства уничтожения.

ПД в течение 3 лет.

Вступили в силу и будут действовать до 1 марта 2029 года Требования к подтверждению уничтожения ПД (утв. Приказом Федеральной службы по надзору в сфере связи, информационных технологий и массовых коммуникаций от 28 октября 2022 г. № 179).

С 1 марта 2023 года операторы при уничтожении ПД должны составить специальный документ – акт об уничтожении персональных данных.

Минимальный срок, в течение которого необходимо хранить акты об уничтожении ПД и выгрузку из журналов регистрации событий, – 3 года.

За невыполнение требований возможно привлечение к ответственности по ч. 5 ст. 13.11 КоАП РФ. – Штраф за нарушение установлен в размере до 90 тыс. р. [16].

Правильное сообщение об утечке персональных данных. Вступил в силу Порядок взаимодействия Роскомнадзора и операторов ПД в рамках ведения реестра учета инцидентов в области ПД (утв. Приказом Федеральной службы по надзору в сфере связи, информационных технологий и массовых коммуникаций от 14 ноября 2022 г. № 187).

Под инцидентом следует понимать любой факт, повлекший неправомерную передачу (предоставление, распространение, доступ) ПД. Операторы обязаны уведомлять Роскомнадзор о таких фактах, повлекших нарушение прав субъектов ПД.

Уведомление может быть первичным (должно быть представлено в течение 24 ч) или дополнительным (предоставляется в течение 72 ч).

Первичное уведомление должно содержать сведения о произошедшем инциденте, его возможных причинах, предполагаемом вреде субъекту ПД, а также возможные меры по устранению вреда. Дополнительное уведомление должно содержать сведения о результатах внутреннего расследования инцидента с указанием виновных лиц в случае, если в ходе расследования они были выявлены.

Новые правила оценки вреда. Вступили в силу и будут действовать до 1 марта 2029 года Требования к оценке вреда, который может быть причинен субъектам ПД в случае нарушения Федерального закона «О персональных данных» (утв. Приказом Федеральной службы по надзору в сфере связи, информационных технологий и массовых коммуникаций от 27 октября 2022 г. № 178).

Оценку вероятного вреда должно осуществлять лицо, ответственное за организацию обработки ПД либо созданная оператором комиссия. Для

целей оценки вреда указанные субъекты определяют одну из степеней вреда, который может быть причинен субъекту ПД в случае нарушения Федерального закона «О персональных данных». Степени разделены на три категории: высокую, среднюю или низкую [28].

Результаты оценки необходимо зафиксировать в акте оценки вреда. Если нарушение будет выявлено в ходе проводимой Роскомнадзором проверки, его нужно будет устранить. Роскомнадзор вправе выдать компании соответствующее предписание.

Следует не пренебрегать возможностью оперативно получать информацию о нововведениях в законодательстве и соответствующих разъяснениях специалистов на сайте Роскомнадзора на соответствующих страницах, например, «Новости», «Персональные данные» и др.

6.3. ОТВЕТСТВЕННОСТЬ ЗА НАРУШЕНИЕ ЗАКОНА О ПЕРСОНАЛЬНЫХ ДАННЫХ

В случае нарушений требований Федерального закона от 27.07.2006 г. № 152-ФЗ «О персональных данных» согласно действующему законодательству предусматриваются различные виды ответственности: дисциплинарная, административная, гражданско-правовая и уголовная ответственность (примеры представлены ниже).

Дисциплинарная ответственность.

Нарушение: Разглашение охраняемой законом тайны (государственной, коммерческой, служебной и иной), ставшей известной работнику в связи с исполнением им трудовых обязанностей, в том числе разглашения персональных данных другого работника.

Норма: Подпункт «в» п. 6 ч. 1 ст. 81 ТК РФ [34].

Санкция: Увольнение.

Нарушение: Иные нарушения в области персональных данных при их обработке.

Норма: Ст. 90 и 191 ТК РФ.

Санкция: Выговор, замечание.

Гражданско-правовая ответственность.

Нарушение: Причинение лицу убытков в результате нарушения правил обработки его персональных данных.

Норма: ст. 15 ГК РФ [7].

Санкция: Возмещение убытков.

Нарушение: Причинение гражданину морального вреда (нравственных страданий) вследствие нарушения правил обработки персональных данных.

Норма: ст. 24 Закона № 152-ФЗ, ст. 151 ГК РФ.

Санкция: Компенсация морального вреда (независимо от возмещения имущественного вреда и понесенных субъектом убытков).

Административная ответственность. Административная ответственность за невыполнение требований законодательства в области обработки персональных данных установлена статьями [16]:

– ст. 13.11 Кодекса Российской Федерации об административных правонарушениях (КоАП РФ) «Нарушение законодательства Российской Федерации в области персональных данных»;

– ст. 19.4 КоАП РФ «Неповиновение законному распоряжению должностного лица органа, осуществляющего государственный надзор (контроль), должностного лица организации, уполномоченной в соответствии с федеральными законами на осуществление государственного надзора, должностного лица органа, осуществляющего муниципальный контроль»;

– ст. 19.4.1 КоАП РФ «Воспрепятствование законной деятельности должностного лица органа государственного контроля (надзора)», должностного лица организации, уполномоченной в соответствии с федеральными законами на осуществление государственного надзора, должностного лица органа муниципального контроля»;

– ст. 19.5 КоАП РФ «Невыполнение в срок законного предписания (постановления, представления, решения)» органа (должностного лица), осуществляющего государственный надзор (контроль), организации, уполномоченной в соответствии с федеральными законами на осуществление государственного надзора (должностного лица), органа (должностного лица), осуществляющего муниципальный контроль»;

– ст. 19.7 КоАП РФ «Непредставление сведений (информации)»;

Дела об административных правонарушениях, предусмотренных ст. 13.11, 13.14, 19.4, 19.4.1, 19.5, 19.7 КоАП РФ, возбуждаются уполномоченным органом и рассматриваются судом.

Уголовная ответственность.

Норма: Ст. 137 (ч.1, 2) УК РФ [36].

Нарушение: Незаконное соби́рание или распространение сведений о частной жизни лица, составляющих его личную или семейную тайну, без его согласия либо распространение этих сведений в публичном выступлении, публично демонстрирующем произведение, или средствах массовой информации.

Спектр отношений, связанных с обработкой персональных данных, достаточно широк. Эти отношения регулируются действующим Федеральным законом от 27.07.2006 № 152-ФЗ «О персональных данных» [24].

При этом действие указанного Федерального закона не распространяется на отношения, возникающие при:

– обработке персональных данных физическими лицами исключительно для личных и семейных нужд, если при этом не нарушаются права субъектов персональных данных;

– организации хранения, комплектования, учета и использования содержащих персональные данные документов Архивного фонда Российской Федерации и других архивных документов в соответствии с законодательством об архивном деле в Российской Федерации;

– обработке персональных данных, отнесенных в установленном порядке к сведениям, составляющим государственную тайну.

Следует отметить, что «предоставление, распространение, передача и получение информации о деятельности судов в Российской Федерации, содержащей персональные данные... осуществляются в соответствии с Федеральным законом от 22 декабря 2008 года № 262-ФЗ «Об обеспечении доступа к информации о деятельности судов в Российской Федерации».

Таким образом, действующие нормативные и правовые акты создают правовые основы защиты персональных данных при использовании современных информационных технологий и систем. Конечно это не «панацея» от существующих и все возрастающих угроз, связанных с оборотом такой специфической информации, как персональные данные, но это именно то, что создает основу системного централизованного подхода к обеспечению защиты персональных данных при неуклонных тенденциях роста использования современных информационных технологий и систем.

ЗАКЛЮЧЕНИЕ

В учебном пособии рассмотрены общие понятия информационных технологий, их классификация, информационные технологии конечного пользователя, технология базы данных: история развития баз данных и их классификация, информация о нейронных сетях, способы информационной безопасности. Подробно изложены принципы работы со справочными правовыми системами.

Надеемся, что материал, предложенный в электронном пособии «Информационные технологии и системы», поможет желающим разобраться в некоторых вопросах информационных технологий и систем.

СПИСОК ИСПОЛЬЗУЕМЫХ ИСТОЧНИКОВ

1. **Агафонов, А. А.** Основы технологий баз данных: учебное пособие / А. А. Агафонов, А. М. Белов. – Самара : Издательство Самарского университета, 2023. – 304 с.
2. **Бурцева, Е. В.** Информационные технологии в юриспруденции: учеб. пособие / Е. В. Бурцева, А. В. Селезнев, В. Н. Чернышов. – Тамбов : Изд-во ФГБОУ ВПО «ТГТУ», 2012. – 104 с.
3. **Гласснер, Э.** Глубокое обучение без математики. Т. 1. Основы / Э. Гласснер ; пер. с английского В. А. Яроцкого. – М. : ДМК Пресс, 2019. – 578 с.
4. **Гласснер, Э.** Глубокое обучение без математики. Т. 2. Практика: руководство / Э. Гласснер ; пер. с англ. В. А. Яроцкого. – М. : ДМК Пресс, 2020. – 610 с.
5. **Гольдберг, Й.** Нейросетевые методы в обработке естественного языка: руководство / Й. Гольдберг ; пер. с англ. А. А. Слинкина. – М. : ДМК Пресс, 2019. – 282 с.
6. **ГОСТ Р 59853–2021.** Информационные технологии (ИТ). Комплекс стандартов на автоматизированные системы. Автоматизированные системы. Термины и определения. – М. : ФГБУ «РСТ», 2021.
7. **Гражданский кодекс Российской Федерации (часть первая)** от 30.11.1994 № 51-ФЗ (ред. от 14.04.2023) // Собрание законодательства Российской Федерации. – 1994. – № 32. – Ст. 3301.
8. **Гудфеллоу, Я.** Глубокое обучение / Я. Гудфеллоу, И. Бенджио, А. Курвилль ; пер. с англ. А. А. Слинкина. – 2-е изд. – М. : ДМК Пресс, 2018. – 652 с.
9. **Дадян, Э. Г.** Методы, модели, средства хранения и обработки данных : учебник / Э. Г. Дадян, Ю. А. Зеленков. – М. : Вузовский учебник: ИНФРА-М, 2022. – 168 с.
10. **Жуковский, О. И.** Информационные технологии в управлении : учеб. пособие / О. И. Жуковский. – Томск : Эль Контент, 2017. – 169 с.
11. **Информатика для юристов и экономистов : учебник для вузов / под ред. С. В. Симоновича.** 2-е изд. – СПб. : Питер, 2014. – 544 с.
12. **Информатика : учебник / под ред. Н. В. Макаровой.** – М. : Финансы и статистика, 2000. – 768 с.
13. **Информационно-правовой портал ГАРАНТ.РУ [сайт].** – URL : <https://www.garant.ru/> (дата обращения: 07.10.2023).
14. **Информационные системы : учеб. пособие / Е. В. Бурцева, И. П. Рак, А. В. Селезнев и др.** – Тамбов : Изд-во Тамб. гос. техн. ун-та, 2009. – 128 с.
15. **Карпова, Т. С.** Базы данных: модели, разработка, реализация : учеб. пособие / Т. С. Карпова – М. : Интуит НОУ, 2016. – 403 с.
16. **Кодекс Российской Федерации об административных правонарушениях** от 30.12.2001 № 195-ФЗ (ред. от 13.06.2023) // Российская газета. – 2001. – № 256. – 31 декабря.

17. **Конституция** Российской Федерации (принята всенародным голосованием 12.12.1993 с изменениями, одобренными в ходе общероссийского голосования 01.07.2020) // Российская газета. – 2020. – № 144. – 4 июля.

18. **КонсультантПлюс** [сайт]. – URL : <https://www.consultant.ru/> (дата обращения : 07.10.2023).

19. **Макарова, Н. В.** Информатика : учебник для вузов / Н. В. Макарова, В. Б. Волков. – СПб. : Питер, 2015. – 576 с.

20. **Минашкин, В. Г.** Статистический анализ использования цифровых технологий в организациях: региональный аспект / В. Г. Минашкин, П. Э. Прохоров // Статистика и экономика. – 2018. – Т. 15, № 5. – С. 51 – 62.

21. **Нестеров, С. А.** Базы данных: учебник и практикум для вузов / С. А. Нестеров. – М. : Издательство Юрайт, 2023. – 230 с.

22. **Новиков, Б. А.** Основы технологий баз данных : учебное пособие / Б. А. Новиков, Е. А. Горшкова, Н. Г. Графеева ; под ред. Е. В. Рогова. – 2-е изд. – М. : ДМК Пресс, 2020. – 582 с.

23. **Об информации**, информационных технологиях и о защите информации: Федеральный закон от 27.07.2006 № 149-ФЗ // Официальный интернет-портал правовой информации: [сайт]. – URL : <http://www.pravo.gov.ru> (дата обращения : 15.08.2023)

24. **О персональных** данных: федеральный закон от 27 июля 2006 г. № 152-ФЗ (ред. 06.02.2023) // Российская газета. – 2006. – № 165. – 29 июля.

25. **Остроух, А. В.** Системы искусственного интеллекта: монография / А. В. Остроух, Н. Е. Суркова. – 2-е изд., стер. – СПб. : Лань, 2021. – 228 с.

26. **Отчет «Digital 2022 Russian Federation»** – Цифровые тенденции в России в 2022 году // Kepios. В: [сайт]. – URL : <https://datareportal.com/reports/digital-2022-russian-federation> (дата обращения : 29.10.2023).

27. **Паттерсон, Д.** Глубокое обучение с точки зрения практика / Д. Паттерсон, А. Гибсон. – М. : ДМК Пресс, 2018. – 418 с.

28. **Петренко, В. И.** Защита персональных данных в информационных системах. Практикум: учебное пособие для СПО / В. И. Петренко, И. В. Мандрица. – 2-е изд., стер. – СПб. : Лань, 2022. – 108 с.

29. **Потопахин, В. В.** Романтика искусственного интеллекта / В. В. Потопахин. – М. : ДМК Пресс, 2017. – 170 с.

30. **Сакова, Т. Г.** Информационные технологии в сфере экономической безопасности : учеб. пособие / Т. Г. Сакова, О. В. Юдина. – Самара : Изд-во Самар. гос. экон. ун-та, 2019. – 105 с.

31. **Система Тоталити** // тоталити: [сайт]. – URL : <https://totality.ru> (дата обращения : 17.09.2023).

32. **Словарь** лингвистических терминов Жеребило // Gufo.me: [сайт]. – URL: <https://gufo.me/search?term=технология> (дата обращения : 14.08.2023).

33. **СЭД «ДЕЛО»** // Электронные Офисные Системы: [сайт]. – URL : https://eos.ru/eos_products/eos_delo/sed-delo/ (дата обращения : 17.09.2023).

34. **Трудовой** кодекс Российской Федерации от 30 декабря 2001 г. № 197-ФЗ (ред. 09.03.2021) // Российская газета. – 2001. – № 256. – 31 декабря.

35. **Тужилин, А. С.** Использование современных информационных технологий / А. С. Тужилин, А. В. Сладков, С. А. Губин // Актуальные исследования. 2023. № 31(161). – С. 10 – 12. – URL : <https://apni.ru/article/6813-ispolzovanie-sovremennikh-it> (дата обращения : 07.11.2023)

36. **Уголовный** кодекс Российской Федерации от 13.06.1996 № 63-ФЗ (ред. от 13.06.2023) // Собрание законодательства РФ. –1996. – № 25. – Ст. 2954.

37. **Фундаментальные** принципы создания автоматизации мест для работы // ARPRIME SOFTWARE. Автоматизация и разработка программ для бизнеса «Арбор Прайм»: [сайт]. – URL : <http://arprime.ru/avtomatizacia/rabocheye-mesto-spetsialista#fundamentalnyye-printsipy-formirovaniya> (дата обращения : 07.11.2023).

38. **Хултен, Д.** Разработка интеллектуальных систем: руководство / Д. Хултен ; пер. с англ. В. С. Яценкова. – М. : ДМК Пресс, 2019. – 284 с.

39. **Чио, К.** Машинное обучение и безопасность: руководство / К. Чио, Д. Фримэн ; пер. с англ. А. В. Снастина. – М. : ДМК Пресс, 2020. – 388 с.

40. **Шалев-Шварц, Ш.** Идеи машинного обучения : учебное пособие / Ш. Шалев-Шварц, Ш. Бен-Давид ; пер. с англ. А. А. Слинкина. – М. : ДМК Пресс, 2019. – 436 с.

ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ	3
1. ВВЕДЕНИЕ В ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ	4
1.1. Понятие информационной технологии	4
1.2. Классификация информационных технологий	7
2. ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ РАБОЧЕГО МЕСТА ПОЛЬЗОВАТЕЛЯ	12
2.1. Автоматизированное рабочее место	12
2.2. Электронный офис	16
3. ТЕХНОЛОГИИ БАЗ ДАННЫХ	21
3.1. Развитие баз данных	21
3.2. Классификация баз данных	24
4. ИСКУССТВЕННЫЕ НЕЙРОННЫЕ СЕТИ	30
4.1. Классификация машинного обучения	30
4.1.1. Контролируемое и неконтролируемое обучение	31
4.1.2. Пакетное и онлайн-обучение.....	33
4.1.3. Обучение на основе примеров или моделей	33
4.2. Проблемы при создании систем машинного обучения	34
4.3. Пять подходов к машинному обучению	35
4.4. Обзор нейронных сетей	40
4.4.1. Размер модели	42
4.4.2. Функции активации	44
4.4.3. Как искусственные нейронные сети «учатся»	46
4.4.4. Оптимизаторы	49
4.4.5. Ранняя остановка, регуляризация и отсев	51
4.5. Вывод	53
5. ИНФОРМАЦИОННЫЕ СИСТЕМЫ И РЕСУРСЫ СЕТИ ИНТЕРНЕТ ДЛЯ ПОИСКА ПРАВОВОЙ ИНФОРМАЦИИ	55
6. ПРАВОВЫЕ ОСНОВЫ ЗАЩИТЫ ПЕРСОНАЛЬНЫХ ДАННЫХ ПРИ ИСПОЛЬЗОВАНИИ СОВРЕМЕННЫХ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ И СИСТЕМ	64
6.1. Нормативные правовые акты в области персональных данных ..	64
6.2. Некоторые особенности обработки персональных данных	68
6.3. Ответственность за нарушение закона о персональных данных .	72
ЗАКЛЮЧЕНИЕ	75
СПИСОК ИСПОЛЬЗУЕМЫХ ИСТОЧНИКОВ	76

Учебное электронное издание

БУРЦЕВА Елена Васильевна
ПЛАТЁНКИН Алексей Владимирович
РАК Игорь Петрович
ТЕРЕХОВ Алексей Васильевич

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ И СИСТЕМЫ

Учебное пособие

Редактирование Е. С. Мордасовой
Графический и мультимедийный дизайнер Н. И. Кужильная
Обложка, упаковка, тиражирование Е. С. Мордасовой

ISBN 978-5-8265-2776-4



Подписано к использованию 16.05.2024.

Тираж 50 шт. Заказ № 64

Издательский центр ФГБОУ ВО «ТГТУ»
392000, г. Тамбов, ул. Советская, д. 106, к. 14
Тел./факс (4752) 63-81-08.
E-mail: izdatelstvo@tstu.ru