ЛЕКЦИЯ 1
Основы и современные направления развития искусственного интеллекта

Цель лекции

Цель лекции — сформировать у обучающихся целостное представление об искусственном интеллекте как научном и технологическом направлении, рассмотреть основные понятия и задачи искусственного интеллекта, этапы его развития, основные подходы к созданию интеллектуальных систем, а также современные направления применения технологий искусственного интеллекта.

Изучение данной темы является основой для последующего освоения машинного и глубокого обучения, обработки естественного языка, компьютерного зрения, генеративного искусственного интеллекта, больших языковых моделей и AI-агентов, которые предусмотрены программой дисциплины.

1. Понятие искусственного интеллекта

Искусственный интеллект является одним из наиболее активно развивающихся направлений современной информатики. Сегодня технологии искусственного интеллекта применяются в медицине, образовании, промышленности, банковской сфере, транспорте, информационной безопасности, электронной коммерции и многих других областях.

При этом понятие «искусственный интеллект» нельзя сводить только к чат-ботам или генеративным системам. Современные большие языковые модели являются лишь одним из направлений искусственного интеллекта.

Искусственный интеллект, Artificial Intelligence, AI, — *это область информатики, занимающаяся разработкой методов, алгоритмов и программных систем, способных выполнять задачи, для решения которых традиционно требуются элементы человеческого интеллекта.*

К таким задачам относятся распознавание объектов, понимание и генерация текста, анализ речи, принятие решений, прогнозирование, обучение на основе данных, планирование действий и поиск решений.

Рассмотрим простой пример. Человек получает электронное письмо и достаточно быстро определяет, является оно обычным сообщением или спамом. Если требуется автоматизировать эту задачу, необходимо создать систему, которая сможет анализировать признаки сообщения и принимать решение о том, к какому классу оно относится.

Такую систему можно считать примером применения искусственного интеллекта.

Другой пример — распознавание лица на фотографии. Человек выполняет эту задачу практически автоматически. Для компьютерной системы изображение первоначально представляет собой набор числовых значений пикселей. Необходимо разработать модель, которая сможет извлечь из этих данных необходимые признаки и определить, какое лицо изображено.

Таким образом, основная идея искусственного интеллекта заключается в создании вычислительных систем, способных выполнять интеллектуальные функции и решать задачи, которые трудно полностью описать набором простых заранее заданных инструкций.

Чем искусственный интеллект отличается от обычной программы?

Для понимания сущности искусственного интеллекта важно сравнить интеллектуальную систему с традиционной программой.

В классическом программировании разработчик заранее определяет правила обработки входных данных.

Например, необходимо написать программу для вычисления стоимости товара со скидкой. Можно задать правило:

если сумма покупки превышает 100 000 тенге, предоставить скидку 10 %.

Здесь программист точно знает условие и самостоятельно записывает его в программном коде.

Схематично традиционное программирование можно представить следующим образом:

Данные + Правила → Результат

Однако существует большое количество задач, для которых невозможно или чрезвычайно сложно заранее описать все правила.

Например, попробуем сформулировать полный набор правил, позволяющий определить, изображена на фотографии кошка или собака. Можно указать форму ушей, размер глаз, шерсть и другие характеристики. Но существует огромное количество пород, ракурсов, условий освещения и вариантов изображения. Создать вручную правила для всех возможных случаев практически невозможно.

Поэтому используется другой подход. Компьютерной системе предоставляется большое количество примеров, на основе которых модель самостоятельно выявляет закономерности.

Например, системе предоставляют тысячи изображений с известными классами «кошка» и «собака». В процессе обучения модель определяет характеристики изображений, которые позволяют различать эти классы.

Такой подход является основой **машинного обучения**.

Интеллектуальная система

С понятием искусственного интеллекта тесно связано понятие интеллектуальной системы.

Интеллектуальная система — *программная или программно-аппаратная система, способная анализировать информацию, использовать знания или данные, принимать решения и выполнять задачи, требующие элементов интеллектуального поведения.*

Интеллектуальная система обычно получает некоторую входную информацию, обрабатывает её с использованием определённых методов или моделей и формирует результат.

Например, интеллектуальная система обнаружения сетевых атак получает данные о сетевом трафике. Система анализирует характеристики соединений и определяет, является ли активность нормальной или потенциально вредоносной.

В медицинской диагностической системе входными данными могут быть результаты анализов, медицинские изображения и сведения о пациенте. Модель анализирует эти данные и формирует прогноз или рекомендацию для специалиста.

Таким образом, общую схему работы интеллектуальной системы можно представить следующим образом:

Входные данные → Обработка/модель ИИ → Решение или прогноз

Однако современная интеллектуальная система часто включает дополнительные этапы: сбор и подготовку данных, обучение модели, оценку её качества и последующее обновление модели при появлении новых данных.

Основные задачи искусственного интеллекта

Искусственный интеллект охватывает множество различных задач.

Одной из наиболее распространённых является **классификация**.

Классификация — задача определения категории или класса, к которому относится исследуемый объект.

Например, система может определять, является банковская операция нормальной или мошеннической. Электронное письмо может классифицироваться как обычное сообщение или спам. Медицинское изображение может быть отнесено к определённой диагностической категории.

Другой важной задачей является **прогнозирование**.

В этом случае система анализирует имеющиеся данные и пытается определить возможное будущее значение или событие.

Например, на основе предыдущих продаж можно прогнозировать спрос на товар. На основании показаний датчиков промышленного оборудования можно оценивать вероятность его отказа.

Искусственный интеллект также используется для **распознавания образов**. Объектами распознавания могут быть изображения, лица, рукописные символы, речь и другие типы информации.

Еще одной задачей является **принятие решений**. Интеллектуальная система анализирует состояние среды и выбирает действие, которое наиболее соответствует поставленной цели.

Например, автономный робот должен определить маршрут движения и избежать препятствий.

Современные системы ИИ также способны выполнять **генерацию контента**: создавать текст, изображения, программный код, аудио и другие виды цифровых данных.

История развития искусственного интеллекта

Идея создания машин, способных имитировать интеллектуальное поведение человека, появилась задолго до современных компьютеров. Однако формирование искусственного интеллекта как самостоятельного научного направления произошло в середине XX века.

Важным этапом стала работа английского математика Алана Тьюринга. В 1950 году он опубликовал статью «Computing Machinery and Intelligence», в которой поставил вопрос о возможности машин демонстрировать интеллектуальное поведение.

С именем Тьюринга связан **тест Тьюринга**. Его основная идея заключается в оценке способности машины вести диалог таким образом, чтобы человек не мог уверенно определить, общается он с человеком или компьютерной системой.

Сам термин **Artificial Intelligence** получил распространение после Дартмутского исследовательского проекта 1956 года, связанного с работами Джона Маккарти и других исследователей.

Первоначально исследователи предполагали, что многие интеллектуальные задачи можно решить достаточно быстро. Разрабатывались программы для решения логических задач, игр и доказательства математических теорем.

Позднее стало понятно, что реальный интеллект значительно сложнее. Компьютерам не хватало вычислительной мощности, объёмов данных и эффективных алгоритмов.

Периоды резкого снижения финансирования и интереса к исследованиям искусственного интеллекта получили название **AI winter — «зима искусственного интеллекта»**.

Новый этап развития начался благодаря увеличению вычислительных возможностей компьютеров, появлению больших массивов цифровых данных и развитию методов машинного обучения.

Особенно заметный прогресс произошёл в 2010-х годах благодаря глубокому обучению, развитию графических процессоров и появлению крупных наборов данных.

В 2020-х годах одним из наиболее заметных направлений стало развитие генеративного искусственного интеллекта и больших языковых моделей.

Таким образом, современный рост возможностей ИИ стал результатом одновременного развития **алгоритмов, данных и вычислительных ресурсов**.

Слабый и общий искусственный интеллект

При обсуждении возможностей искусственного интеллекта важно различать специализированный и общий интеллект.

Узкий или слабый искусственный интеллект, Artificial Narrow Intelligence, ANI, — *система ИИ, предназначенная для решения определённой задачи или ограниченного класса задач.*

Практически все реально применяемые сегодня системы искусственного интеллекта относятся именно к этому типу.

Например, система распознавания лиц хорошо решает задачу идентификации по изображению, но сама по себе не умеет управлять автомобилем или прогнозировать погоду.

Система рекомендаций интернет-магазина предлагает пользователю товары, но не выполняет медицинскую диагностику.

Другим понятием является **Artificial General Intelligence, AGI — общий искусственный интеллект**.

Под AGI обычно понимается гипотетическая интеллектуальная система, способная выполнять широкий спектр интеллектуальных задач и переносить полученные знания между различными областями.

Важно понимать, что AGI является концепцией и исследовательской целью, а не общепринятым обозначением уже достигнутого уровня современных систем.

Поэтому нельзя автоматически считать любую мощную языковую модель общим искусственным интеллектом.

Машинное обучение как направление искусственного интеллекта

Одним из важнейших направлений современного искусственного интеллекта является **машинное обучение, Machine Learning, ML**.

Машинное обучение — *направление искусственного интеллекта, в котором алгоритмы строят модели на основе данных и используют выявленные закономерности для получения прогнозов или принятия решений.*

Главное отличие машинного обучения от традиционного программирования заключается в том, что все правила не задаются разработчиком вручную.

Рассмотрим задачу обнаружения мошеннических банковских операций.

В традиционной программе можно установить правило: если сумма операции превышает определённое значение и операция выполняется из другой страны, считать её подозрительной.

Однако мошеннические схемы постоянно меняются. Одного или нескольких правил может быть недостаточно.

При машинном обучении используется набор предыдущих операций. Для каждой операции известны её характеристики и результат: нормальная она или мошенническая. Алгоритм анализирует эти примеры и строит модель, способную классифицировать новые операции.

Таким образом, данные становятся одним из основных ресурсов интеллектуальной системы.

В вашем силлабусе машинному обучению посвящён отдельный блок курса: рассматриваются обучение с учителем, без учителя и с подкреплением, а также классификация, регрессия, кластеризация и снижение размерности.

Глубокое обучение и нейронные сети

Нейронная сеть — *вычислительная модель, состоящая из взаимосвязанных вычислительных элементов и способная изменять свои параметры в процессе обучения на данных.*

Современные нейронные сети могут содержать большое количество слоёв.

Глубокое обучение, Deep Learning, DL, — *направление машинного обучения, основанное на использовании многослойных нейронных сетей для автоматического выявления сложных закономерностей в данных.*

Особенно эффективно глубокое обучение применяется при обработке изображений, речи и естественного языка.

Например, при распознавании изображения нейронная сеть может постепенно формировать различные уровни представления информации. На ранних этапах она выявляет простые характеристики изображения, а на более глубоких уровнях — более сложные структуры, необходимые для классификации объекта.

Связь между рассматриваемыми понятиями можно условно представить следующим образом:

Искусственный интеллект → Машинное обучение → Глубокое обучение

То есть глубокое обучение является частью машинного обучения, а машинное обучение — одним из основных направлений искусственного интеллекта.

Обработка естественного языка

Обработка естественного языка, Natural Language Processing, NLP, — *направление искусственного интеллекта, связанное с компьютерной обработкой, анализом, пониманием и генерацией человеческого языка.*

К задачам NLP относятся классификация текстов, машинный перевод, анализ тональности, извлечение информации, ответы на вопросы, автоматическое реферирование и генерация текста.

Например, компании ежедневно получают тысячи отзывов клиентов. Человеку сложно вручную проанализировать такое количество сообщений. NLP-модель может автоматически определить, какие отзывы являются положительными, отрицательными или нейтральными.

Другой пример — автоматический анализ юридических документов. Система может находить в большом массиве документов необходимые сущности, даты, организации или определённые смысловые фрагменты.

Развитие NLP непосредственно связано с современными большими языковыми моделями, которые будут отдельно рассматриваться в последующих лекциях курса. Согласно силлабусу, обработка естественного языка изучается на 9-й неделе, а архитектура и применение больших языковых моделей — на 12-й.

Компьютерное зрение

Компьютерное зрение, Computer Vision, — направление искусственного интеллекта, разрабатывающее методы получения и анализа информации из изображений и видео.

Человек, посмотрев на фотографию улицы, легко определяет автомобили, людей, здания и дорожные знаки. Для компьютера изображение первоначально представляет собой массив числовых значений.

Задача компьютерного зрения заключается в том, чтобы преобразовать эти значения в содержательную информацию.

Например, интеллектуальная система видеонаблюдения может обнаруживать людей и транспортные средства. Медицинская система может анализировать рентгеновские или другие диагностические изображения. Система контроля качества на производстве может автоматически обнаруживать дефекты продукции.

Компьютерное зрение стало одним из направлений, где развитие глубоких нейронных сетей привело к существенному повышению качества решения практических задач.

Генеративный искусственный интеллект

Традиционные модели искусственного интеллекта часто предназначены для анализа существующих данных. Например, классификатор получает изображение и определяет его класс.

Генеративный искусственный интеллект, Generative AI, — *направление искусственного интеллекта, в котором модели способны создавать новый контент на основе закономерностей, изученных по обучающим данным.*

Генерируемым контентом может быть текст, изображение, программный код, аудио, видео и другие виды данных.

Например, языковая модель может создать краткое содержание документа, подготовить проект письма, объяснить программный код или сформировать текст по заданной теме.

Генеративная модель изображений может создать новое изображение по текстовому описанию.

При этом важно понимать, что генеративная модель не является обычной базой данных, из которой просто извлекается заранее подготовленный ответ. Она формирует новый результат на основе параметров и закономерностей, сформированных в процессе обучения.

Генеративный ИИ является отдельным крупным блоком вашего курса: в силлабусе ему посвящены 11-я лекция, а далее изучаются LLM, Prompt Engineering и AI-агенты.

Большие языковые модели

Одним из наиболее известных современных направлений генеративного ИИ являются **большие языковые модели — Large Language Models, LLM**.

Большая языковая модель — *модель искусственного интеллекта, обученная на больших объёмах текстовых и других данных и способная работать с последовательностями естественного языка, включая генерацию и преобразование текста.*

Современные языковые модели могут отвечать на вопросы, объяснять сложные понятия, переводить текст, анализировать документы, помогать при программировании и выполнять другие задачи.

При взаимодействии с такими моделями важную роль играет запрос пользователя — **prompt**.

Отсюда развивается отдельное направление — **Prompt Engineering**, то есть методы проектирования и оптимизации запросов для получения более подходящих результатов от моделей.

Однако результаты LLM необходимо оценивать критически. Модель может сформировать убедительно звучащий, но фактически неправильный ответ. Поэтому использование генеративного ИИ не отменяет необходимость проверки полученной информации.

AI-агенты и Agentic AI

Следующим этапом развития интеллектуальных систем является переход от моделей, которые только отвечают на запрос, к системам, способным выполнять последовательность действий для достижения поставленной цели.

AI-агент — *программная интеллектуальная система, которая получает информацию о задаче или среде, принимает решения и выполняет действия для достижения определённой цели.*

Например, обычная языковая модель может объяснить, как составить аналитический отчёт.

AI-агент потенциально может выполнить более сложную последовательность: получить необходимые данные из разрешённых источников, обработать их, провести анализ, сформировать таблицу и подготовить итоговый документ.

С этим связано понятие **Agentic AI** — подход к созданию систем ИИ, обладающих большей автономностью при планировании и выполнении многошаговых задач в рамках заданных ограничений.

Если несколько агентов взаимодействуют друг с другом для решения общей задачи, формируется **многоагентная система**.

Например, один агент может заниматься поиском и структурированием информации, другой — анализом данных, третий — проверкой результата.

В вашем курсе AI-агенты и многоагентные интеллектуальные системы выделены в отдельную тему 14-й недели.

Искусственный интеллект в информационной безопасности

Поскольку дисциплина предназначена для магистрантов образовательной программы «Комплексное обеспечение информационной безопасности», особое значение имеет применение искусственного интеллекта в сфере кибербезопасности. Это соответствует и заявленным результатам курса, где предусмотрены разработка и применение моделей ИИ для задач безопасности, а также анализ эффективности, надёжности и безопасности интеллектуальных моделей.

Одним из примеров является **обнаружение аномалий**.

В информационной системе формируется большое количество событий: входы пользователей, сетевые соединения, обращения к файлам, действия приложений. Модель может изучать характеристики нормальной активности и выявлять необычные события.

Другим направлением является анализ сетевого трафика. Модели машинного обучения могут классифицировать сетевые соединения и обнаруживать признаки потенциальных атак.

ИИ также применяется при анализе вредоносного программного обеспечения, обнаружении фишинговых сообщений, анализе журналов событий и поддержке работы специалистов центров мониторинга безопасности.

Однако искусственный интеллект одновременно создаёт новые риски. Генеративные модели могут использоваться злоумышленниками для масштабирования социальной инженерии и создания вводящего в заблуждение контента. Кроме того, сами системы искусственного интеллекта могут становиться объектом атак.

Поэтому специалист по информационной безопасности должен рассматривать ИИ одновременно **как инструмент защиты и как новый объект обеспечения безопасности**.

Этика и ответственное применение искусственного интеллекта

Развитие искусственного интеллекта связано не только с техническими возможностями, но и с социальными, этическими и правовыми вопросами.

Ответственный искусственный интеллект, Responsible AI, — *подход к разработке и использованию систем ИИ с учётом безопасности, справедливости, прозрачности, конфиденциальности и ответственности за принимаемые решения.*

Например, если система ИИ используется для принятия решений, затрагивающих людей, необходимо понимать, какие данные использовались, насколько корректно работает модель и не создаёт ли она систематического неравного отношения к отдельным группам.

Отдельная проблема связана с персональными данными. Для обучения и работы моделей могут использоваться большие массивы информации, поэтому необходимо учитывать требования к конфиденциальности и защите данных.

Также возникают вопросы авторского права, достоверности сгенерированного контента и ответственности за решения, принятые с использованием ИИ.

Именно поэтому в вашем силлабусе этика и безопасность не вынесены за рамки технической подготовки. Курс предусматривает формирование навыков безопасного и ответственного использования ИИ с учётом этических требований, авторского права и защиты информации.

Современная структура искусственного интеллекта

Рассмотренные направления позволяют увидеть, что искусственный интеллект представляет собой не одну технологию, а целую совокупность взаимосвязанных областей.

Условно эту структуру можно представить следующим образом:

- Искусственный интеллект (AI)
- Машинное обучение (ML)
- Глубокое обучение (DL)
- Обработка естественного языка (NLP)
- Компьютерное зрение (Computer Vision)
- Генеративный ИИ (Generative AI)
- Большие языковые модели (LLM)
- AI-агенты и многоагентные системы

При этом такая схема не означает, что все направления строго вложены друг в друга. Например, NLP и Computer Vision являются самостоятельными областями, в которых могут использоваться методы машинного и глубокого обучения. Генеративный ИИ также может работать с текстом, изображениями, аудио и другими типами данных.

Поэтому современный искусственный интеллект правильнее рассматривать как систему взаимосвязанных методов, моделей и технологий.

Заключение

Искусственный интеллект представляет собой область информатики, направленную на создание методов и программных систем, способных выполнять задачи, связанные с анализом информации, обучением, распознаванием, прогнозированием, генерацией данных и принятием решений.

Современный искусственный интеллект развивается благодаря сочетанию трёх ключевых факторов: совершенствованию алгоритмов, росту объёмов доступных данных и увеличению вычислительных ресурсов.

Одним из основных направлений ИИ является машинное обучение, позволяющее строить модели на основе данных. Его дальнейшее развитие привело к широкому применению глубоких нейронных сетей. Эти методы лежат в основе современных технологий компьютерного зрения, обработки естественного языка и многих других интеллектуальных систем.

В последние годы особое значение приобрели генеративный искусственный интеллект, большие языковые модели, Prompt Engineering, AI-агенты и многоагентные системы. Именно поэтому эти направления включены в последующие темы курса наряду с классическими методами машинного обучения.

При этом развитие возможностей искусственного интеллекта повышает требования к его безопасному и ответственному применению. Специалист должен не только уметь использовать модель, но и понимать её ограничения, оценивать качество результатов, учитывать вопросы информационной безопасности, конфиденциальности, этики и ответственности.

Контрольные вопросы

1. Что понимается под искусственным интеллектом?
2. Чем система искусственного интеллекта отличается от традиционной программы?
3. Что называется интеллектуальной системой?
4. Какие задачи могут решаться методами искусственного интеллекта?
5. Какую роль сыграли работы Алана Тьюринга в развитии ИИ?
6. Что означает понятие «зима искусственного интеллекта»?
7. Чем узкий искусственный интеллект отличается от концепции AGI?
8. Что называется машинным обучением?
9. Чем машинное обучение отличается от традиционного программирования?

10. Что такое глубокое обучение?
11. Как связаны AI, Machine Learning и Deep Learning?
12. Какие задачи решает обработка естественного языка?
13. Что понимается под компьютерным зрением?
14. Что такое генеративный искусственный интеллект?
15. Что называется большой языковой моделью?
16. Что такое AI-агент?
17. Для каких задач искусственный интеллект может применяться в информационной безопасности?
18. Какие риски связаны с использованием технологий ИИ?
19. Что понимается под ответственным применением искусственного интеллекта?
20. Какие направления развития искусственного интеллекта вы считаете наиболее перспективными и почему?