Минобрнауки России

Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Оренбургский государственный университет»

Кафедра компьютерной безопасности и математического обеспечения информационных систем

РАБОЧАЯ ПРОГРАММА

ДИСЦИПЛИНЫ

«Б1.Д.В.Э.3.1 Основы искусственного интеллекта»

Уровень высшего образования

СПЕЦИАЛИТЕТ

Специальность
10.05.01 Компьютерная безопасность
(код и наименование специальности)

специализация №3 «Разработка защищенного программного обеспечения»
(наименование направленности (профиля)/специализации образовательной программы)

Квалификация
Специалист по защите информации

Форма обучения
Очная

Год набора 2021

Рабочая программа дисциплины «*Б1.Д.В.Э.3.1 Основы искусственного интеллекта*» рассмотрена и утверждена на заседании кафедры

Кафедра компьютерной безопасности и математического обеспечения информационных систем
наименование кафедры

протокол № 10 от "21" 06 2021 г.

Заведующий кафедрой
Кафедра компьютерной безопасности и математического обеспечения информационных систем
наименование кафедры *подпись* И.В. Влацкая *расшифровка подписи*

Исполнители:
ст. преподаватель *должность* *подпись* Н.А. Заельская *расшифровка подписи*

должность *подпись* *расшифровка подписи*

СОГЛАСОВАНО:
Председатель методической комиссии по специальности
10.05.01 Компьютерная безопасность
код наименование *личная подпись* *расшифровка подписи*

Заведующий отделом комплектования научной библиотеки
личная подпись Н.Н. Бигалиева *расшифровка подписи*

Уполномоченный по качеству факультета
личная подпись И.В. Крючкова *расшифровка подписи*

№ регистрации _____

© Заельская Н.А., 2021
© ОГУ, 2021

1 Цели и задачи освоения дисциплины

Цель (цели) освоения дисциплины:

изучить и уметь применять на практике основные понятия и методы искусственного интеллекта и интеллектуальных систем.

Задачи:

- рассмотреть историю становления и развития искусственного интеллекта;
- рассмотреть технические постановки основных задач, решаемых системами искусственного интеллекта;
- ознакомить с современными областями исследования по искусственному интеллекту;
- ознакомить с основными моделями представления знаний и некоторыми интеллектуальными системами;
- рассмотреть теоретические и некоторые практические вопросы создания и эксплуатации экспертных систем.

2 Место дисциплины в структуре образовательной программы

Дисциплина относится к дисциплинам (модулям) по выбору вариативной части блока Д «Дисциплины (модули)»

Пререквизиты дисциплины: *Б1.Д.Б.13 Информатика*

Постреквизиты дисциплины: *Отсутствуют*

3 Требования к результатам обучения по дисциплине

Процесс изучения дисциплины направлен на формирование следующих результатов обучения

Код и наименование формируемых компетенций	Код и наименование индикатора достижения компетенции	Планируемые результаты обучения по дисциплине, характеризующие этапы формирования компетенций
УК-1 Способен осуществлять критический анализ проблемных ситуаций на основе системного подхода, вырабатывать стратегию действий	УК-1-В-2 Осуществляет критический анализ и синтез информации, полученной из разных источников УК-1-В-4 Применяет методы сбора, хранения, обработки, передачи, анализа и синтеза информации с использованием компьютерных технологий для решения поставленных задач	**Знать:** – понятия инженерии знаний и нейрокибернетики; – методы представления и обработки знаний; – структуру экспертных систем и их архитектурные особенности в зависимости от особенностей решаемой задачи; – этапы построения экспертных систем; – основные модели нейронных сетей, методов и алгоритмов их обучения. **Уметь:** – ориентироваться в различных типах интеллектуальных систем; – ориентироваться в различных методах представления знаний, переходить от одного метода к другому; – формализовать знания экспертов с применением различных методов представления знаний. **Владеть:** – навыками постановки задачи построения экспертной системы для решения задачи выбора вариантов в плохо формализуемой предметной области;

Код и наименование формируемых компетенций	Код и наименование индикатора достижения компетенции	Планируемые результаты обучения по дисциплине, характеризующие этапы формирования компетенций
		– методами разработки продукционных баз знаний для решения задач выбора вариантов в плохо формализуемой предметной области.

4 Структура и содержание дисциплины

4.1 Структура дисциплины

Общая трудоемкость дисциплины составляет 4 зачетные единицы (144 академических часа).

Вид работы	Трудоемкость, академических часов	
	10 семестр	всего
Общая трудоёмкость	**144**	**144**
Контактная работа:	**74,25**	**74,25**
Лекции (Л)	30	30
Лабораторные работы (ЛР)	44	44
Промежуточная аттестация (зачет, экзамен)	0,25	0,25
Самостоятельная работа: *- самоподготовка (проработка и повторение лекционного материала и материала учебников и учебных пособий;* *- подготовка к лабораторным занятиям;* *- подготовка к рубежному контролю и т.п.)*	**69,75**	**69,75**
Вид итогового контроля (зачет, экзамен, дифференцированный зачет)	**зачет**	

Разделы дисциплины, изучаемые в 10 семестре

№ раздела	Наименование разделов	Количество часов				
		всего	аудиторная работа			внеауд. работа
			Л	ПЗ	ЛР	
1	Технологии ИИ в современном мире	18	4	-	6	8
2	Экспертные системы	32	4	-	16	12
3	Модели поиска решений в ЭС	34	6	-	16	12
4	Нейронные сети	18	4	-	6	8
5	Большие языковые модели	14	4	-	-	10
6	Популярные реализации технологии машинного обучения.	14	4	-	-	10
7	Кодекс этики и Национальная стратегия ИИ	14	4	-	-	10
	Итого:	144	30	-	44	70
	Всего:	144	30	-	44	70

4.2 Содержание разделов дисциплины

1 Технологии ИИ в современном мире

Основные этапы развития ИИ.

Лабиринтная гипотеза: мышление как поиск пути в пространстве состояний. Машинное обучение: автоматическое приобретение знаний (с учителем, без учителя, с подкреплением).

Метаобучение: «обучение обучению», адаптация к новым задачам по малым данным. Воплощённые системы (Embodied AI): ИИ с физическим/виртуальным телом, обучающийся через взаимодействие со средой. Слабый и сильный ИИ: различие между узкоспециализированными системами и гипотетическим универсальным интеллектом (AGI).

История и текущая ситуация.

Этапы: от экспертных систем (1950-е) → машинное обучение → современные мультимодальные модели. Лидеры рынка: США и Китай; роль России и Национальная стратегия развития ИИ до 2030 г.

Определение и технологии ИИ.

ИИ как комплекс технологий, имитирующих когнитивные функции человека. Ключевые свойства: понимание языка, обучение, мышление, действие.

Направления исследований.

По задачам: машинный перевод, компьютерное зрение, речевые системы и др.. По инструментам: нейросети, эволюционные вычисления, экспертные системы. По моделям мышления: поиск решений, представление знаний, машинное обучение.

Теоретические основы.

Два подхода: «интеллектуальный агент» (экспертные системы) и «нейронные сети». Сравнение достоинств и недостатков каждого подхода. Перспективы синтеза методов для практического применения.

2 Экспертные системы и модели поиска решений в ЭС

Структура и назначение ЭС.

Определение: программная система, имитирующая действия эксперта в узкой предметной области. Исторический контекст: переход от универсальных программ к специализированным системам на основе знаний (конец 1970-х).

Архитектура типовой ЭС: База знаний (БЗн) + Механизм логического вывода (МЛВ) + Рабочая память (РП) = СОЗ. Лингвистический процессор (интерфейс). Компонент приобретения знаний (КПЗн). Компонент объяснений (КОб) — ключевое отличие ЭС от других систем.

Преимущества: решение неформализуемых задач, диалоговый режим, тиражирование экспертизы.

Классификация и этапы разработки.

Классификация по задачам: интерпретация, прогноз, диагностика, проектирование, планирование, обучение. По типу знаний: поверхностные (правила) и глубинные (модели, причинно-следственные связи). По времени: статические и динамические (реальное время). По интеграции: автономные и гибридные

Этапы разработки: Идентификация (задачи, цели, участники, ресурсы). Концептуализация (ключевые понятия, отношения, стратегии). Формализация (выбор формализма, моделирование). Выполнение (прототипирование, итеративное развитие). Тестирование (верификация БЗ и механизмов вывода). Опытная эксплуатация (оценка полезности и удобства).Методы извлечения знаний: наблюдение, обсуждение, анализ «вслух», проверка прототипа.

Представление знаний.

Свойства знаний (6 ключевых): Внутренняя интерпретация (метаданные). Внутренняя структура связей. Внешняя (ситуативная) структура связей. Возможность шкалирования. Семантическая метрика (близость понятий). Активность (способность порождать новые знания). Проблемы семантики: омонимия (одна форма — много смыслов) и синонимия (много форм — один смысл)

Классификация знаний: факты/эвристики; декларативные/процедурные; интенсиональные/экстенсиональные; предметные/прагматические. Модели представления: логические, продукционные, семантические сети, фреймы, гибридные.

Продукционная модель представлений знаний

Принцип работы: знания представляются в виде правил вида «ЕСЛИ <условие>, ТО <действие>». Структура продукции: антецедент (условие) → консеквент (действие) → постусловие. Типы ядер правил: детерминированные (однозначные) и недетерминированные (вероятностные, нечёткие)

Стратегии логического вывода: Прямой вывод (*forward chaining*) — от данных к цели. Обратный вывод (*backward chaining*) — от гипотезы к фактам. Смешанный вывод. Практическая реализация: использование стека целей, алгоритм *Rete* для оптимизации сопоставления.

Семантические сети.

Принцип: представление знаний в виде ориентированного графа (узлы = понятия/объекты, дуги = отношения). Типы отношений: лингвистические, логические, теоретико-множественные (ISA, HAS-PART, AKO), квантифицированные.

Механизмы вывода: Наследование свойств по иерархии. Сопоставление по совпадению (*pattern matching*).

Особенности: поддержка транзитивности, экономия памяти за счёт наследования, динамические сети для сценариев.

Фреймовая модель.

Принцип: фрейм — структурированный шаблон для описания стереотипных ситуаций/объектов. Структура: имя фрейма + слоты (имя, значение, тип) + присоединённые процедуры. Типы слотов и процедур: Демоны: IF-ADDED, IF-REMOVED, IF-NEEDED, IF-REPLACED. Процедуры-слуги для вычисления значений.

Механизм вывода: наследование по AKO-связям, согласование, детализация, обновление фреймов.

Применение: экспертные системы, NLP, робототехника

3 Нейронные сети

Искусственный нейрон: понятие и принцип работы.

Биологический прототип: нейрон мозга состоит из тела, дендритов (входы) и аксона (выход); передача сигнала через синапсы с помощью нейротрансмиттеров. Математическая модель Маккалока–Питтса (1943): искусственный нейрон как вычислительный элемент с тремя компонентами: Синапсы с весами (w_i) — усиление/ослабление входных сигналов. Сумматор — вычисление взвешенной суммы: $S = \Sigma(x_i \cdot w_i) + w_0$. Функция активации — нелинейное преобразование: $Y = f(S)$. Ограничения: одиночный нейрон может моделировать только простейшие логические операции (И, ИЛИ, НЕ)

Нейронная сеть (*Neural Network*).

Определение: математическая модель, построенная по принципу организации биологических нейронных сетей; состоит из связанных процессоров-нейронов. Ключевое преимущество: возможность обучения (настройки весов), а не программирования в традиционном смысле.

Классификация ИНС. По топологии: полносвязные, многослойные (слоистые). По обучению: с учителем, без учителя. По типу связей: прямого распространения, рекуррентные. По типу сигнала: бинарные, аналоговые.

Однослойный персептрон: простейшая модель с жёсткой пороговой функцией; ограничен в возможностях. Многослойный персептрон (MLP): наличие скрытых слоёв позволяет решать сложные нелинейные задачи. Алгоритм обучения (дельта-правило): итеративная корректировка весов на основе ошибки между желаемым и фактическим выходом.

Функции активации.

Пороговая функция. Линейная функция. ReLU (*Rectified Linear Unit*). Сигмоида. Гиперболический тангенс.

Выбор функции: зависит от задачи; ReLU — для скрытых слоёв, сигмоида/тангенс — для классификации, линейная — для регрессии на выходе.

LLM: большие языковые модели.

Определение: передовые системы ИИ на базе глубокого обучения, способные понимать, обрабатывать и генерировать человекоподобный текст; обучаются на гигантских массивах данных, содержат миллиарды параметров. Техническая основа — архитектура *Transformer*: Механизм самовнимания (*self-attention*): взвешивание важности всех слов в контексте при обработке каждого токена. Параллельная обработка последовательностей (в отличие от *RNN*). Архитектура энкодер-декодер для анализа и генерации. Позиционное кодирование для учёта порядка элементов. Масштаб и обучение: «Большими» называют из-за миллиардов/триллионов параметров. Два этапа: предобучение на общих текстах → тонкая настройка (*fine-tuning*) под конкретные задачи. Мультимодальность: современные *LLM* работают не только с текстом, но и с изображениями, аудио, видео

Задачи NLP: понимание смысла, анализ структур, генерация ответов. Базовые технологии: Токенизация — разбиение текста на минимальные единицы. Синтаксический и семантический анализ. *NER* (распознавание именованных сущностей). Распознавание намерений (*intent recognition*). Анализ тональности, машинный перевод. Голосовые системы: *ASR* (распознавание речи) + *TTS* (синтез речи) + диалоговый *AI*. Понимание контекста: запоминание истории диалога, адаптация под пользователя.

Возможности и ограничения генерации текстов.Типы генерируемого контента: академические работы, художественные тексты, реклама, деловая документация, контент для соцсетей. Настраиваемые параметры: объём, стиль, целевая аудитория, контекст использования. Режим рассуждения: демонстрация логической цепочки, анализ данных, обоснование выводов — повышает прозрачность и обучаемость. Ограничения: отсутствие доступа к интернету в реальном времени, риск «галлюцинаций», отсутствие сознания и этической ответственности.

Генерация изображений и промпт-инжиниринг. Принцип: создание визуального контента по текстовому описанию (промпту). Структура эффективного промпта: <основное> + <детали> + <стиль>. Рекомендации: чёткая формулировка роли, цели, пошаговых инструкций, формата вывода и ограничений. Практика: итеративное улучшение запросов, критическая оценка результатов.

4 Популярные реализации технологии машинного обучения

Ключевые направления применения машинного обучения. Компьютерное зрение. Биометрическая идентификация. Обработка естественного языка (*NLP*). Распознавание речи. Синтез речи (TTS). Машинное зрение (промышленное). Машинный перевод. Генерация текстов. Диалоговые системы (чат-боты). Анализ тональности.

5 Кодекс этики и Национальная стратегия ИИ

Принципы этики и правила поведения. Главный приоритет — защита прав и интересов человека. Осознание ответственности при создании и использовании ИИ. Ответственность за последствия применения ИИ несёт человек. Применение ИИ по назначению и во благо людям. Прозрачность и правдивость в информировании. Правовая основа: Конституция РФ, национальные стратегии, международные договоры. Добровольный характер присоединения; назначение уполномоченных по этике в организациях. Создание Комиссии по реализации Кодекса и ведение реестра участников.

Цели, задачи и меры развития ИИ в РФ для обеспечения национальных интересов и технологического суверенитета. Правовая основа: Конституция РФ, ФЗ «Об информации», «О персональных данных», указы Президента о стратегиях развития. Ключевые понятия: ИИ, технологии ИИ, большие генеративные и фундаментальные модели, доверенные технологии, параметры модели, отказоустойчивость.

ИИ — драйвер экономического роста, инноваций и повышения качества жизни. Вызовы для РФ: нехватка вычислительных мощностей, дефицит кадров, нормативные барьеры, риски информационной безопасности, ограничения доступа к технологиям.

Основные принципы развития и использования технологий ИИ. Защита прав и свобод человека. Безопасность: недопустимость умышленного вреда, минимизация рисков. Прозрачность: объяснимость алгоритмов, недискриминационный доступ к информации. Технологический суверенитет: приоритет отечественных решений и программно-аппаратных комплексов. Целостность инновационного цикла: связь науки и реального сектора. Эффективное использование: приоритет действующих механизмов господдержки. Поддержка конкуренции: недопустимость монополизации. Открытость и доступность: ограничение доступа только в сферах госуправления и ОПК. Преемственность: постепенный переход органов власти к использованию ИИ. Защищённость: безопасность, правовая охрана, разграничение ответственности. Достоверность исходных данных: методологическое обеспечение качества данных.

4.3 Лабораторные работы

№ ЛР	№ раздела	Наименование лабораторных работ	Кол-во часов
1	1	Правильная постановка целей.	2

№ ЛР	№ раздела	Наименование лабораторных работ	Кол-во часов
2	1	Реализация модели предметной области поставленной задачи.	4
3	3	Элементы формальной логики для управления техническими устройствами.	4
4	3	Реализация продукционной модели предметной области.	4
5	3	Построение семантической сети проблемной ситуации.	4
6	3	Создание фрейма и системы фреймов.	4
7	2	Методы практического извлечения знаний.	4
8	2	Проектирование экспертной системы.	6
9	2	Реализация разработанной экспертной системы.	6
10	4	Реализация алгоритма обучения нейронной сети Хопфилда и/или Хэмминга.	6
		Итого:	44

5 Учебно-методическое обеспечение дисциплины

5.1 Основная литература

1 Сотник, С. Л. Проектирование систем искусственного интеллекта: курс / С. Л. Сотник. – Москва : Интернет-Университет Информационных Технологий (ИНТУИТ), 2007. – 204 с. : ил., табл., схем. – Режим доступа: по подписке. – URL: https://biblioclub.ru/index.php?page=book&id=234802

2 Веретехина, С. В. Модели, методы, алгоритмы и программные решения вычислительных машин, комплексов и систем : учебник : [16+] / С. В. Веретехина, В. Л. Симонов, О. Л. Мнацаканян. – Изд. 2-е, доп. – Москва ; Берлин : Директ-Медиа, 2021. – 307 с. : ил., схем., табл. – Режим доступа: по подписке. – URL: https://biblioclub.ru/index.php?page=book&id=602526

5.2 Дополнительная литература

1 Сергеев, Н. Е. Системы искусственного интеллекта : учебное пособие : [16+] / Н. Е. Сергеев. – Таганрог : Южный федеральный университет, 2016. – Ч. 1. – 123 с. : схем., ил., табл. – Режим доступа: по подписке. – URL: https://biblioclub.ru/index.php?page=book&id=493307

2 Павлов, С. И. Системы искусственного интеллекта : учебное пособие / С. И. Павлов. – Томск : Томский государственный университет систем управления и радиоэлектроники, 2011. – Ч. 1. – 175 с. – Режим доступа: по подписке. – URL: https://biblioclub.ru/index.php?page=book&id=208933

3 Павлов, С. И. Системы искусственного интеллекта : учебное пособие / С. И. Павлов. – Томск : Томский государственный университет систем управления и радиоэлектроники, 2011. – Ч. 2. – 194 с. – Режим доступа: по подписке. – URL: https://biblioclub.ru/index.php?page=book&id=208939

5.3 Периодические издания

1 Вестник компьютерных и информационных технологий : журнал. - М. : Агентство "Роспечать", 2021.

2 Информационные технологии : журнал. - М. : Агентство "Роспечать", 2021.

5.4 Интернет-ресурсы

www.citforum.ru/ - портал аналитических и научных статей в области информационных технологий;

www.intuit.ru/ - Национальный открытый университет «ИНТУИТ»

5.5 Программное обеспечение, профессиональные базы данных и информационные справочные системы

1 Операционная система Microsoft Windows текущей версии. Доступна в рамках подписки Microsoft DreamSpark Premium. Разработчик: компания Microsoft. Режим доступа: https://e5.onthehub.com/WebStore/ProductsByMajorVersionList.aspx?cmi_mnuMain=bdba23cf-e05e-e011-971f-0030487d8897&ws=58727022-4bac-e211-88b7-f04da23e67f4&vsro=8

2 Офисный пакет Microsoft Office (Word, Excel, Power Point) текущей версии. Доступен в рамках лицензионного соглашения OVS-ES. Разработчик: компания Microsoft. Режим доступа: https://products.office.com/en/home

6 Материально-техническое обеспечение дисциплины

Учебные аудитории для проведения занятий лекционного типа, для проведения групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, оснащены комплектами ученической мебели, мультимедийным проектором, доской и экраном.

Для проведения лабораторных занятий используется компьютерный класс, оснащенный комплектами ученической мебели, доской и компьютерами с обеспечением доступа в электронную информационно-образовательную среду ОГУ.

Помещение для самостоятельной работы обучающихся оснащено компьютерной техникой, с обеспечением доступа в электронную информационно-образовательную среду ОГУ.