

Министерство науки и высшего образования Российской Федерации
ФГБОУ ВО «Уральский государственный лесотехнический университет»
Социально-экономический институт
Кафедра интеллектуальных систем

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

включая фонд оценочных средств и методические указания
для самостоятельной работы обучающихся

Б1.В.01 – ИНТЕЛЛЕКТУАЛЬНЫЕ МЕТОДЫ ИЗВЛЕЧЕНИЯ ЗНАНИЙ

Направление подготовки – 09.04.02 Информационные системы и технологии

Направленность (профиль) – Интеллектуальные информационные системы и технологии

Квалификация – магистр

Количество зачётных единиц (часов) – 3 (108)

г. Екатеринбург, 2025

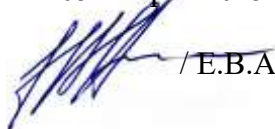
Разработчики: к.с.-х.н., доцент



/Е.В.Анянова/;

Рабочая программа утверждена на заседании кафедры интеллектуальных систем (протокол № 9 от «05» марта 2025 года).

Зав. Кафедрой



/Е.В.Анянова /

Рабочая программа рекомендована к использованию в учебном процессе методической комиссией социально-экономического института (протокол № 2 от «06» марта 2025 года).

Председатель методической комиссии СЭИ



/ А.В. Чевардин /

Рабочая программа утверждена директором социально-экономического института

Директор СЭИ



/Ю.А. Капустина/

«06» марта 2025 года.

Оглавление

1. Общие положения	4
2. Перечень планируемых результатов обучения по дисциплине, соотнесенных с планируемыми результатами освоения образовательной программы	4
3. Место дисциплины в структуре образовательной программы	5
4. Объем дисциплины в зачетных единицах с указанием количества академических часов, выделенных на контактную работу обучающихся с преподавателем (по видам учебных занятий) и на самостоятельную работу обучающихся	6
5. Содержание дисциплины, структурированное по темам (разделам) с указанием отведенного на них количества академических часов	6
5.1. <i>Трудоемкость разделов дисциплины</i>	6
5.2. <i>Содержание занятий лекционного типа</i>	6
5.3 <i>Темы и формы практических занятий</i>	8
5.4 <i>Самостоятельная работа обучающихся</i>	8
6. Перечень учебно-методического обеспечения по дисциплине	9
7. Фонд оценочных средств для проведения промежуточной аттестации обучающихся по дисциплине	11
7.1. <i>Перечень компетенций с указанием этапов их формирования в процессе освоения образовательной программы</i>	11
7.2. <i>Описание показателей и критериев оценивания компетенций на различных этапах их формирования, описание шкал оценивания</i>	11
7.3. <i>Типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций в процессе освоения образовательной программы</i>	12
7.4. <i>Соответствие шкалы оценок и уровней сформированных компетенций</i>	14
8. Методические указания для самостоятельной работы обучающихся	14
9. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине	15
10.Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине	16

1. Общие положения

Дисциплина «Интеллектуальные методы извлечения знаний» относится к дисциплинам части, формируемой участниками образовательных отношений, блока Б1 «Дисциплины (модули)» учебного плана, входящего в состав образовательной программы высшего образования 09.04.02 Информационные системы и технологии (профиль «Интеллектуальные информационные системы и технологии»).

Нормативно-методической базой для разработки рабочей программы учебной дисциплины «Интеллектуальные методы извлечения знаний» являются:

- Федеральный закон «Об образовании в Российской Федерации», утвержденный приказом Минобрнауки РФ № 273-ФЗ от 29.12.2012;
- Приказ Минобрнауки России № 245 от 06.04.2021 г. Об утверждении порядка организации и осуществления образовательной деятельности по образовательным программам высшего образования – программам бакалавриата, программам специалитета, программам магистратуры;
- Приказ Министерства труда и социальной защиты от 17.09.2014 г. №645н «Об утверждении профессионального стандарта «Руководитель разработки программного обеспечения»;
- Приказ Министерства труда и социальной защиты от 17.09.2014 г. №647н «Об утверждении профессионального стандарта «Администратор баз данных»;
- Приказ Министерства труда и социальной защиты от 18.11.2014 г. №893н «Об утверждении профессионального стандарта «Руководитель проектов в области информационных технологий»;
- Федеральный государственный образовательный стандарт высшего образования – магистратура по направлению подготовки 09.04.02 «Информационные системы и технологии», утвержденный приказом Министерства образования и науки РФ № 917 от 19.09.2017 (ФГОС ВО);
- Учебный план образовательной программы высшего образования направления 09.04.02 «Информационные системы и технологии» (профиль «Интеллектуальные информационные системы и технологии») подготовки магистров по очной форме обучения, одобренный Ученым советом УГЛУ (протокол №3 от 20.03.2025).

Обучение по образовательной программе 09.04.02 «Информационные системы и технологии (профиль «Интеллектуальные информационные системы и технологии») осуществляется на русском языке.

2. Перечень планируемых результатов обучения по дисциплине, соотнесенных с планируемыми результатами освоения образовательной программы

Планируемыми результатами обучения по дисциплине, являются знания, умения, владения и/или опыт деятельности, характеризующие этапы/уровни формирования компетенций и обеспечивающие достижение планируемых результатов освоения образовательной программы в целом.

Цель изучения дисциплины – изучение средств и методов интеллектуального анализа данных для разработки интеллектуальных систем.

Задачи дисциплины:

- сформировать знания об алгоритмах интеллектуального анализа данных;
- сформировать умения использовать инструменты интеллектуального анализа данных;
- сформировать навыки применения в практической деятельности средств и методов интеллектуального анализа данных.

Процесс изучения дисциплины направлен на формирование следующих компетенций:

– **ПК-2** – способен проводить анализ системных проблем обработки информации на уровне базы данных, внедрять в практику новые технологии работы с базой данных, обеспечивающей эффективное функционирование интеллектуальных информационных систем в организации.

Достижение планируемых результатов освоения образовательной программы обеспечивается через получение результатов обучения по дисциплинам и практикам.

В результате освоения дисциплины обучающийся должен:

знать:

– основные типы задач, решаемых методами интеллектуального извлечения знаний, в рамках автоматизации деятельности человека в различных отраслях экономики, связанных в том числе с принятием решений, управлением технологическими процессами, распознаванием образов, ситуаций и процессов;

– алгоритмы интеллектуальной обработки данных, их достоинства и недостатки, области применения;

– инструментальные средства реализации алгоритмов интеллектуального извлечения знаний;

уметь:

– выявлять проблемы организации, связанные с обработкой информации на уровне базы данных; формулировать проблему в виде задачи интеллектуального извлечения знаний;

– определять класс задач, относящихся к интеллектуальным, применять известные способы интеллектуальной обработки данных, способы представления знаний к практической задаче;

– применять алгоритмы и их ансамбли для интеллектуального анализа данных;

– осуществлять подбор инструментов, реализующих алгоритмы интеллектуального анализа данных, и применять их для решения поставленной задачи;

владеть навыками:

– анализа системных проблем обработки информации и постановки проблемы в виде задачи интеллектуального анализа данных;

– применения методов и алгоритмов интеллектуальной обработки данных, инструментов, реализующих алгоритмы;

– анализа качества применяемых алгоритмов обработки информации на уровне базы данных и возможностей их внедрения в практику работы организации.

3. Место дисциплины в структуре образовательной программы

Данная учебная дисциплина относится к дисциплинам части, формируемой участниками образовательных отношений, блока Б1 «Дисциплины (модули)» учебного плана, что означает формирование в процессе обучения у обучающихся профессиональных компетенций в рамках выбранного профиля подготовки.

Освоение данной дисциплины является необходимой основой для последующего изучения дисциплин ОПОП и написания выпускной квалификационной работы.

Перечень обеспечивающих, сопутствующих и обеспечиваемых дисциплин

Обеспечивающие	Сопутствующие	Обеспечиваемые
	Программирование в среде Python Основы нейронных сетей	Методы анализа больших наборов данных / Экспертные системы Нейронные сети в TensorFlow Производственная практика (преддипломная) Выполнение и защита выпускной квалификационной работы

Указанные связи дисциплины дают обучающемуся системное представление о комплексе изучаемых дисциплин в соответствии с ФГОС ВО, что обеспечивает требуемый

теоретический уровень и практическую направленность в системе обучения и будущей деятельности выпускника.

4. Объем дисциплины в зачетных единицах с указанием количества академических часов, выделенных на контактную работу обучающихся с преподавателем (по видам учебных занятий) и на самостоятельную работу обучающихся

Общая трудоемкость дисциплины

Вид учебной работы	Всего академических часов
Контактная работа с преподавателем*:	52,25
лекции (Л)	18
практические занятия (ПЗ)	-
лабораторные работы (ЛР)	34
Промежуточная аттестация (ПА)	0,25
Самостоятельная работа обучающихся:	55,75
изучение теоретического курса	29
подготовка к текущему контролю	20
курсовая работа (курсовой проект)	-
подготовка к промежуточной аттестации	6,75
Вид промежуточной аттестации:	зачет
Общая трудоемкость, з.е./ часы	3 / 108

*Контактная работа обучающихся с преподавателем, в том числе с применением дистанционных образовательных технологий, включает занятия лекционного типа, и (или) занятия семинарского типа, лабораторные занятия, и (или) групповые консультации, и (или) индивидуальную работу обучающегося с преподавателем, а также аттестационные испытания промежуточной аттестации. Контактная работа может включать иные виды учебной деятельности, предусматривающие групповую и индивидуальную работу обучающихся с преподавателем. Часы контактной работы определяются Положением об организации и проведении контактной работы при реализации образовательных программ высшего образования, утвержденным Ученым советом УГЛУТУ от 25 февраля 2020 года.

5. Содержание дисциплины, структурированное по темам (разделам) с указанием отведенного на них количества академических часов

5.1. Трудоемкость разделов дисциплины

№ п/п	Наименование раздела дисциплины	Л	ПЗ	ЛР	Всего контактной работы	Самостоятельная работа
1	Раздел 1. Основы интеллектуального анализа данных	2	-	-	2	5
2	Раздел 2. Введение в машинное обучение	2	-	4	6	5
3	Раздел 3. Задачи классификации	2	-	4	6	6
4	Раздел 4. Задачи прогнозирования	2	-	4	6	6
5	Раздел 5. Кластеризация	2	-	4	6	6
6	Раздел 6. Задачи понижения размерности	2	-	4	6	6
7	Раздел 7. Поиск ассоциаций	2	-	4	6	5
8	Раздел 8. Решающие деревья	2	-	4	6	5
9	Раздел 9. Ансамблирование методов извлечения знаний	2	-	6	8	5
Итого по разделам:		18	-	34	52	49
Промежуточная аттестация		-	-	-	0,25	6,75
Всего		108				

5.2. Содержание занятий лекционного типа

Раздел 1. Основы интеллектуального анализа данных

Понятие Data Mining. Предпосылки развития автоматических методов анализа данных. Методы извлечения знаний и области их применения. Основные этапы эволюции ал-

горитмов машинного обучения. Предметно-ориентированные аналитические системы. Статистические пакеты. Нейронные сети. Рекомендательные системы. Деревья решений (decision trees). Системы для визуализации многомерных данных. Инструменты для анализа и извлечения данных

Раздел 2. Введение в машинное обучение

Понятие машинного обучения. Модель машинного обучения. Классы задач, решаемых методами машинного обучения. Задачи классификации. Прогнозные задачи. Задачи кластеризации. Задачи сжатия данных (уменьшения размерности). Задачи выявления аномалий.

Технология решения задач машинного обучения: постановка задачи, подготовка данных, определение обучающей и тестовой выборки, подбор алгоритма, обучение, определение качества алгоритма, интерпретация результатов.

Данные для машинного обучения, их релевантность, когерентность, точность, достаточность, полнота, актуальность. Работа с данными перед машинным обучением: сбор данных, нормирование данных, приведение к начальной форме, очистка данных, кодирование данных. Качество данных. Определение объектов и признаков. Признаки для разных типов данных. Разметка данных. Методы и алгоритмы извлечения признаков. Целевая переменная. Функция активации. Функция потерь. Методы оптимизации функции потерь. Обучающая выборка. Тестовая выборка.

Алгоритм обучения. Алгоритм предсказания. Критерии для отбора. Проверка правильности алгоритма. Потери при обучении. Качество алгоритма. Переобучение.

Техники машинного обучения: обучение с учителем (supervised learning), обучение без учителя (unsupervised learning), обучение с частичным привлечением учителя (semi-supervised learning), обучение с подкреплением (reinforcement learning).

Раздел 3. Задачи классификации

Классификация как обучение с учителем. Постановка задачи, ее особенности. Примеры задач. Бинарная и мультикатегорийная классификация. Скалирование. Сбалансированная классификация. Алгоритмы классификации. Методы построения правил классификации. Линейные классификаторы с разделяющей гиперплоскостью в качестве границы. Логистическая регрессия: суть задачи, область применимости, достоинства и недостатки. Метод градиентного спуска. Стохастический градиентный спуск. Метод опорных векторов (алгоритм SVM): суть задачи, область применимости, достоинства и недостатки. Плоскость решений. Метод Байеса. Наивный байесовский алгоритм. Метод ближайших соседей.

Метрики качества классификации: accuracy, точность (Precision), полнота (Recall), f-мера, кривая ROC, кривая Precision-Recall, AUC-ROC, AUC-PR. Инструменты для решения задач классификации.

Раздел 4. Задачи прогнозирования

Задача регрессии как обучение с учителем. Примеры задач на регрессию. Линейная регрессия: суть задачи, область применимости, достоинства и недостатки. Мультиколлинеарность. Регуляризация. Метрики качества регрессии: MAE, MSE, RMSE, Quantile loss, MAPE. Градиентный спуск. Инструменты для решения задач регрессии

Раздел 5. Кластеризация

Постановка задачи кластеризации как обучение без учителя. Алгоритмы кластеризации. Иерархические алгоритмы. Итеративные алгоритмы. Плотностные алгоритмы. Модельные алгоритмы. Визуализация кластеризации. Метод kNN: суть задачи, область применимости, достоинства и недостатки. Метод DBSCAN. Иерархическая кластеризация. Измерение качества кластеризации. Инструменты для решения задач кластеризации

Раздел 6. Задачи понижения размерности

Задача понижения размерности как обучение без учителя. Области применения: сжатие данных, ускорение предсказаний, визуализация данных, описание данных. Алгоритмы понижения размерности: отбор признаков, выделение новых признаков на основе исходных.

Алгоритмы отбора признаков: метод фильтрации, оберточные методы отбора признаков, встроенные методы отбора признаков. Преимущества методов отбора признаков.

Алгоритмы выделения признаков: метод главных компонент, автокодировщик. Алгоритмы визуализации данных: многомерное шкалирование (MDS), метод t-SNE. Применение визуализации.

Понижение размерности в работе с текстами. Измерение качества алгоритмов: восстановление старых признаков, использование метрики той задачи, в которой применяются новые признаки (классификации, регрессии). Инструменты для решения задач понижения размерности

Раздел 7. Поиск ассоциаций

Постановка задачи поиска ассоциаций как обучение без учителя. Правила ассоциации. Параметры отбора ассоциаций: поддержка, достоверность (доверие). Этапы решения задач на поиск ассоциаций. Алгоритмы поиска ассоциативных правил: Apriori, FP-Growth. Инструменты для поиска ассоциативных правил.

Раздел 8. Решающие деревья

Понятие дерева. Решающее дерево. Типы задач, решаемых с помощью деревьев. Задача классификации. Алгоритм построения дерева решений. Правила остановки построения дерева. Обзор алгоритмов, реализующих деревья решений. Обучение деревьев. Обучающие данные. Метод подбора гиперпараметров. Ограничения на деревьях (гиперпараметры). Оценка качества обучения.

Раздел 9. Ансамблирование методов извлечения знаний

Объединение алгоритмов для повышения качества решения задач. Виды ансамблей: бэггинг, стэкинг, бустинг. Комбинирование ансамблей. Random Forest: преимущества и недостатки. Взвешенные ансамбли. Применение ансамблирования в задачах предсказания, классификации. Преимущества и недостатки ансамблевых моделей. Инструменты для обучения ансамблей.

5.3 Темы и формы практических занятий

Учебным планом по дисциплине предусмотрены лабораторные занятия.

№	Наименование раздела дисциплины (модуля)	Форма проведения занятия	Трудоемкость, ч
2	Раздел 2. Введение в машинное обучение Подготовка данных к использованию в машинном обучении. Извлечение признаков. Разметка данных. Определение обучающей и тестовой выборки.	Лабораторная работа	4
3	Раздел 3. Задачи классификации Линейная классификация	Лабораторная работа	4
4	Раздел 4. Задачи прогнозирования Линейная регрессия. Логистическая регрессия	Лабораторная работа	4
5	Раздел 5. Кластеризация Кластеризация. Иерархическая кластеризация. Визуализация кластеризации	Лабораторная работа	4
6	Раздел 6. Задачи понижения размерности Метод фильтрации. Оберточные методы отбора признаков. Метод главных компонент. Визуализация данных.	Лабораторная работа	4
7	Раздел 7. Поиск ассоциаций	Лабораторная работа	4
8	Раздел 8. Решающие деревья	Лабораторная работа	4

	Задача классификации. Случайный лес.		
9	Раздел 9. Ансамблирование интеллектуальных методов извлечения знаний Бэггинг. Стэкинг. Бустинг	Лабораторная работа	6
Итого часов:			34

5.4 Детализация самостоятельной работы

№	Наименование раздела дисциплины (модуля)	Вид самостоятельной работы	Трудоемкость, ч
1	Раздел 1. Основы интеллектуального анализа данных	Изучение рекомендованной литературы, выполнение домашних заданий, подготовка рефератов	5
2	Раздел 2. Введение в машинное обучение	Изучение рекомендованной литературы, выполнение домашних заданий, подготовка рефератов	5
3	Раздел 3. Задачи классификации	Изучение рекомендованной литературы, выполнение домашних заданий, подготовка рефератов	6
4	Раздел 4. Задачи прогнозирования	Изучение рекомендованной литературы, выполнение домашних заданий, подготовка рефератов	6
5	Раздел 5. Кластеризация	Изучение рекомендованной литературы, выполнение домашних заданий, подготовка рефератов	6
6	Раздел 6. Задачи понижения размерности	Изучение рекомендованной литературы, выполнение домашних заданий, подготовка рефератов	6
7	Раздел 7. Поиск ассоциаций	Изучение рекомендованной литературы, выполнение домашних заданий, подготовка рефератов	5
8	Раздел 8. Решающие деревья	Изучение рекомендованной литературы, выполнение домашних заданий, подготовка рефератов	5
9	Раздел 9. Ансамблирование методов извлечения знаний	Изучение рекомендованной литературы, выполнение домашних заданий, подготовка рефератов	5
10	Промежуточная аттестация	Подготовка ответов на вопросы к зачету	6,75
Итого:			55,75

6. Перечень учебно-методического обеспечения по дисциплине **Основная и дополнительная литература**

№	Автор, наименование	Год издания	Примечание
Основная литература			
1	Диане, С. А. Методы искусственного интеллекта: Практикум : учебное пособие / С. А. Диане, А. Д. Воронков. — Москва : РТУ МИРЭА, 2025. — 68 с. — ISBN 978-5-7339-2483-0. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/493502 . — Режим доступа: для авториз. пользователей.	2025	Полнотекстовый доступ при входе по логину и паролю*
2	Методы искусственного интеллекта : учебное пособие / Н. В. Андреянов, Т. С. Евдокимова, А. Д. Павлов [и др.]. — Казань : КНИТУ-КАИ, 2024. — 392 с. — ISBN 978-5-7579-2709-1. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/434162 . — Режим доступа: для авториз. пользователей.	2024	Полнотекстовый доступ при входе по логину и паролю*
3	Манько, С. В. Методы искусственного интеллекта : методические указания / С. В. Манько, С. А. Диане. — Москва : РТУ МИРЭА, 2023. — 38 с. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/382400 . — Режим доступа: для авториз. пользователей.	2023	Полнотекстовый доступ при входе по логину и паролю*
4	Газанова, Н. Ш. Методы искусственного интеллекта : учебно-методическое пособие / Н. Ш. Газанова, С. Н. Никольский. — Москва : РТУ МИРЭА, 2023. — 102 с. — ISBN 978-5-7339-1805-	2023	Полнотекстовый доступ при входе по логину и паролю*

№	Автор, наименование	Год издания	Примечание
	1. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/368756 . — Режим доступа: для авториз. пользователей.		
5	Басаргин, А. А. Методы искусственного интеллекта : учебное пособие / А. А. Басаргин. — Новосибирск : СГУГиТ, 2022. — 164 с. — ISBN 978-5-907513-45-7. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/317474 . — Режим доступа: для авториз. пользователей.	2022	Полнотекстовый доступ при входе по логину и паролю*
6	Пальмов, С. В. Методические указания к лабораторным работам по дисциплине «Интеллектуальный анализ данных» для студентов очной формы обучения направления «Информационные системы и технологии» : методические указания / С. В. Пальмов. — Самара : ПГУТИ, 2018. — 74 с. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/182296 . — Режим доступа: для авториз. пользователей.	2018	Полнотекстовый доступ при входе по логину и паролю*
Дополнительная литература			
7	Келлехер, Д. Наука о данных: базовый курс: [16+] / Д. Келлехер, Б. Тирни; науч. ред. З. Мамедьяров; пер. с англ. М. Белоголовского. — Москва: Альпина Паблишер, 2020. — 224 с.: схем., табл. — Режим доступа: по подписке. — URL: https://biblioclub.ru/index.php?page=book&id=598235 . — ISBN 978-5-9614-3170-4. — Текст: электронный.	2020	Полнотекстовый доступ при входе по логину и паролю*
8	Макшанов, А. В. Технологии интеллектуального анализа данных: учебное пособие / А. В. Макшанов, А. Е. Журавлев. — 2-е изд., стер. — Санкт-Петербург: Лань, 2019. — 212 с. — ISBN 978-5-8114-4493-9. — Текст: электронный // Лань: электронно-библиотечная система. — URL: https://e.lanbook.com/book/120063 . — Режим доступа: для авториз. пользователей.	2019	Полнотекстовый доступ при входе по логину и паролю*
9	Головко, В. А. Нейросетевые технологии обработки данных : учебное пособие / В. А. Головко, В. В. Краснопрошин. — Минск : БГУ, 2017. — 263 с. — ISBN 978-985-566-467-4. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/180542 . — Режим доступа: для авториз. пользователей.	2017	Полнотекстовый доступ при входе по логину и паролю*
10	Воронина, В. В. Теория и практика машинного обучения: учебное пособие / В. В. Воронина. — Ульяновск: УлГТУ, 2017. — 290 с. — ISBN 978-5-9795-1712-4. — Текст: электронный // Лань: электронно-библиотечная система. — URL: https://e.lanbook.com/book/165053 . — Режим доступа: для авториз. пользователей.	2017	Полнотекстовый доступ при входе по логину и паролю*

* прежде чем пройти по ссылке, необходимо войти в систему.

Функционирование электронной информационно-образовательной среды обеспечивается соответствующими средствами информационно-коммуникационных технологий.

Электронные библиотечные системы

Каждый обучающийся обеспечен доступом к электронной библиотечной системе УГЛУ (<http://lib.usfeu.ru/>), ЭБС Издательства Лань <http://e.lanbook.com/>, ЭБС Университетская библиотека онлайн <http://biblioclub.ru/>, содержащих издания по основным изучаемым дисциплинам и сформированных по согласованию с правообладателями учебной и учебно- методической литературы.

Справочные и информационные системы

1. Справочно-правовая система «Консультант Плюс». – Режим доступа: для авториз. пользователей.
2. Информационно-правовой портал Гарант. – URL: <http://www.garant.ru/> – Режим доступа: свободный.
3. База данных Scopus компании ElsevierB.V. – URL: <https://www.scopus.com/> – Режим доступа: для авториз. пользователей.

Профессиональные базы данных

1. Научная электронная библиотека eLibrary. – URL: <http://elibrary.ru/>. – Режим доступа: свободный.
2. Федеральная служба государственной статистики. Официальная статистика. – URL: <http://www.gks.ru/>. – Режим доступа: свободный.
3. Официальный интернет-портал правовой информации. – URL: <http://pravo.gov.ru/> – Режим доступа: свободный.
4. Российская государственная библиотека. – URL: <http://www.rsl.ru/>. – Режим доступа: свободный.
5. Портал искусственного интеллекта. – URL: <http://www.aiportal.ru/>. – Режим доступа: свободный.
6. Хабр. Сообщество ИТ-специалистов. – URL: <https://habr.com/ru/>. – Режим доступа: свободный.
7. Smart Market: Ассистенты Салют, технологии и сервисы Сбера для решения любых ИТ-задач [Open Source]. – URL: <https://developers.sber.ru/>. – Режим доступа: свободный.

Прочие интернет-ресурсы

1. Воронцов, К. Школа Анализа Данных (Яндекс): Машинное обучение / К. Воронцов // Национальный Открытый Университет «Интуит». – URL: <https://intuit.ru/studies/courses/13844/1241/info>. — Режим доступа: свободный.
2. Аникин, Ю. Академия BIG DATA: Введение в аналитику больших массивов данных / Ю. Аникин, И. Борисова, Н. Загоруйко, А. Зырянов, Е. Павловский // Национальный Открытый Университет «Интуит». – URL: <https://intuit.ru/studies/courses/12385/1181/info>. — Режим доступа: свободный.

7. Фонд оценочных средств для проведения промежуточной аттестации обучающихся по дисциплине

7.1. Перечень компетенций с указанием этапов их формирования в процессе освоения образовательной программы

Формируемые компетенции	Вид и форма контроля
ПК-2 – способен проводить анализ системных проблем обработки информации на уровне базы данных, внедрять в практику новые технологии работы с базой данных, обеспечивающей эффективное функционирование интеллектуальных информационных систем в организации	Промежуточный контроль: контрольные вопросы к зачету Текущий контроль: домашние задания, рефераты

7.2 Описание показателей и критериев оценивания компетенций на различных этапах их формирования, описание шкал оценивания

Критерии оценивания устного ответа на контрольные вопросы зачета (промежуточный контроль формирования компетенции ПК-2):

«зачтено» – обучающийся на высоком уровне и в полной мере владеет знаниями основных типов задач, решаемых методами интеллектуального извлечения знаний, алгоритмов интеллектуальной обработки данных, их достоинств и недостатков, инструментальные средства для реализации рассмотренных алгоритмов; самостоятельно и на высоком уровне умеет выявлять проблемы организации, связанные с обработкой

информации на уровне базы данных, и формулировать ее в виде задачи интеллектуального извлечения знаний, применять известные способы интеллектуальной обработки данных, способы представления знаний к практической задаче, осуществлять подбор инструментов, реализующих алгоритмы интеллектуального анализа данных, и применять их для решения поставленной задачи; на высоком уровне владеет навыками анализа проблем обработки информации и постановки проблемы в виде задачи интеллектуального анализа данных, применения методов и алгоритмов интеллектуальной обработки данных, а также инструментов, реализующих их;

«зачтено» – обучающийся практически без ошибок демонстрирует владение знаниями основных типов задач, решаемых методами интеллектуального извлечения знаний, алгоритмов интеллектуальной обработки данных, их достоинств и недостатков, инструментальных средства для реализации рассмотренных алгоритмов; при небольшой коррекции действий умеет выявлять проблемы организации, связанные с обработкой информации на уровне базы данных, и формулировать ее в виде задачи интеллектуального извлечения знаний, применять известные способы интеллектуальной обработки данных, способы представления знаний к практической задаче, осуществлять подбор инструментов, реализующих алгоритмы интеллектуального анализа данных, и применять их для решения поставленной задачи; на хорошем уровне владеет навыками анализа проблем обработки информации и постановки проблемы в виде задачи интеллектуального анализа данных, применения методов и алгоритмов интеллектуальной обработки данных, а также инструментов, реализующих их;

«зачтено» – обучающийся с ошибками, но в объеме, достаточном для дальнейшей профессиональной деятельности, демонстрирует владение знаниями основных типов задач, решаемых методами интеллектуального извлечения знаний, алгоритмов интеллектуальной обработки данных, их достоинств и недостатков, инструментальных средства для реализации рассмотренных алгоритмов; под руководством умеет выявлять проблемы организации, связанные с обработкой информации на уровне базы данных, и формулировать ее в виде задачи интеллектуального извлечения знаний, применять известные способы интеллектуальной обработки данных, способы представления знаний к практической задаче, осуществлять подбор инструментов, реализующих алгоритмы интеллектуального анализа данных, и применять их для решения поставленной задачи; на базовом уровне владеет навыками анализа проблем обработки информации и постановки проблемы в виде задачи интеллектуального анализа данных, применения методов и алгоритмов интеллектуальной обработки данных, а также инструментов, реализующих их;

«не зачтено» – обучающийся не владеет либо владеет фрагментарными знаниями основных типов задач, решаемых методами интеллектуального извлечения знаний, алгоритмов интеллектуальной обработки данных, их достоинств и недостатков, инструментальных средства для реализации рассмотренных алгоритмов; не умеет либо практически не умеет выявлять проблемы организации, связанные с обработкой информации на уровне базы данных, и формулировать ее в виде задачи интеллектуального извлечения знаний, применять известные способы интеллектуальной обработки данных, способы представления знаний к практической задаче, осуществлять подбор инструментов, реализующих алгоритмы интеллектуального анализа данных, и применять их для решения поставленной задачи; не владеет либо демонстрирует очень слабый уровень владения навыками анализа проблем обработки информации и постановки проблемы в виде задачи интеллектуального анализа данных, применения методов и алгоритмов интеллектуальной обработки данных, а также инструментов, реализующих их.

Критерии оценивания рефератов (текущий контроль формирования компетенции ПК-2):

«отлично»: работа выполнена в соответствии с требованиями, выбранная тема раскрыта полностью, материал актуален и достаточен, обучающийся четко и без ошибок

ответил на все контрольные вопросы;

«хорошо»: работа выполнена в соответствии с требованиями, выбранная тема раскрыта, материал актуален, обучающийся ответил на все контрольные вопросы с замечаниями;

«удовлетворительно»: работа выполнена в соответствии с требованиями, выбранная тема частично раскрыта, по актуальности доклада есть замечания, обучающийся ответил на все контрольные вопросы с замечаниями;

«неудовлетворительно»: обучающийся не подготовил работу или подготовил работу, не отвечающую требованиям, ответил на контрольные вопросы с ошибками или не ответил на большую часть вопросов.

Критерии оценивания домашних заданий (текущий контроль формирования компетенции ПК-2)

Оценка	Критерии оценивания
«5» (отлично)	Задачи решены на 8 и более баллов
«4» (хорошо)	Задачи решены на 6-7 баллов
«3» (удовлетворительно)	Задачи решены на 4-5 баллов
«2» (неудовлетворительно)	Задачи решены на 0-3 балла

7.3. Типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций в процессе освоения образовательной программы

Вопросы к зачету (промежуточный контроль)

1. Постановка задачи машинного обучения. Подготовка данных.
2. Типы обучения: с учителем, без учителя, с подкреплением, с частичным участием учителя, активное обучение.
3. Пороговые условия. Precision-Recall и ROC кривые. AUC.
4. Ошибка внутри и вне выборки. Ошибка обобщения.
5. Метрические классификаторы. kNN.
6. Байесовский классификатор.
7. Линейная регрессия.
8. Логистическая регрессия.
9. Метод опорных векторов. Постановка задачи. Формулировка и решение двойственной задачи. Типы опорных векторов. Ядра.
10. Кластеризация. kMeans, DBSCAN.
11. Деревья решений. Информационный выигрыш, критерий Джини. Регуляризация деревьев.
12. Ансамблевые методы.
13. Бэггинг. Bagging.
14. Бустинг. AdaBoost.
15. Случайные леса.
16. Бустинг деревьев решений.
17. Градиентный спуск.

Пример домашнего задания (текущий контроль)

Домашнее задание №1

Домашнее задание №1 выдается студентам в одном варианте и состоит из 3 задач. Каждой задаче присвоен свой балл. Срок выполнения домашнего задания – 2 недели. Форма представления обучающимися домашнего задания – представленные в письменном виде решения задач.

Пример домашнего задания №1:

Задача 1 [3 балла]. Реализуйте алгоритм kNN классификации по k ближайшим соседям, используя простое евклидовое расстояние.

Задача 2 [3 балла]. Реализуйте алгоритм k-means для кластеризации на 2-4 кластера.

Задача 3 [4 балла]. Реализуйте алгоритм DBSCAN, найдите параметры для кластеризации на 4 кластера.

Домашнее задание №2

Домашнее задание №2 выдается студентам в одном варианте и состоит из 4 задач. Каждой задаче присвоен свой балл. Срок выполнения домашнего задания – 2 недели. Форма представления обучающимися домашнего задания – представленные в письменном виде решения задач.

Пример домашнего задания №2:

Задача 1 [3 балла]. Реализуйте алгоритмы построения дерева с критерием информационного выигрыша и критерием Джини и определению класса по мажоритарному классу в листе. Найдите оптимальную глубину дерева в обоих случаях (в отрезке 2-10).

Задача 2 [3 балла]. Примените метод SVM (например, из библиотеки sklearn) для датасета blobs2. Визуализируйте результат (разбиение плоскости и опорные вектора) при разных вариантах ядер (линейное; полиномиальное степеней 2,3,5; RBF).

Задача 3 [4 балла]. Спрогнозируйте среднюю (медианную) стоимость домов в районах Сызрани для датасета sizr_data, используя линейную регрессию. Найдите в таблице с весами признак, соответствующий проценту низкообеспеченного населения в районе. Если значение этого процента велико, будет ли предсказываться высокая средняя стоимость домов или низкая? Логичен ли такой результат?

Примерные темы рефератов (текущий контроль)

1. Нечеткий алгоритм кластеризации c-means.
2. Классификация: метод опорных векторов.
3. Обучение с учителем: ранжирование
4. Random Forest: преимущества и недостатки.
5. Градиентный бустинг.
6. Логистическая регрессия.
7. Логистические методы классификации.
8. Методология CRISP-DM
9. Байесовская классификация
10. Алгоритмы отбора эталонов
11. Критерий информационного выигрыша и критерий Джини
12. Линейный дискриминантный анализ
13. Голосование
14. Бутстраппинг
15. Адаптивный бустинг
16. Градиентный спуск
17. Метод DBSCAN
18. Метод сдвига среднего значения (MeanShift)

7.4. Соответствие шкалы оценок и уровней сформированных компетенций

Уровень сформированности компетенций	Оценка	Пояснения
Высокий	«Зачтено»	Теоретическое содержание курса освоено полностью, обучающийся способен самостоятельно проводить анализ системных проблем обработки информации на уровне базы данных, осуществлять поиск и внедрять в практику новые технологии работы с базой данных, обеспечивающей эффективное функционирование интеллектуальных информационных

		систем в организации, все предусмотренные программой обучения учебные задания выполнены
Базовый	«Зачтено»	Теоретическое содержание курса освоено полностью, обучающийся способен проводить анализ системных проблем обработки информации на уровне базы данных, осуществлять поиск и внедрять в практику новые технологии работы с базой данных, обеспечивающей эффективное функционирование интеллектуальных информационных систем в организации, все предусмотренные программой обучения учебные задания выполнены с незначительными замечаниями.
Пороговый	«Зачтено»	Теоретическое содержание курса освоено полностью, обучающийся способен под руководством проводить анализ системных проблем обработки информации на уровне базы данных, осуществлять поиск и внедрять в практику новые технологии работы с базой данных, обеспечивающей эффективное функционирование интеллектуальных информационных систем в организации, большинство предусмотренных программой обучения учебных заданий выполнено, в них имеются ошибки.
Низкий	«Не зачтено»	Теоретическое содержание курса не освоено, обучающийся не способен проводить анализ системных проблем обработки информации на уровне базы данных, осуществлять поиск и внедрять в практику новые технологии работы с базой данных, обеспечивающей эффективное функционирование интеллектуальных информационных систем в организации, большинство предусмотренных программой обучения учебных заданий либо не выполнены, либо содержат грубые ошибки; дополнительная самостоятельная работа над материалом не привела к какому-либо значительному повышению качества выполнения учебных заданий.

8. Методические указания для самостоятельной работы обучающихся

При самостоятельной проработке курса обучающиеся должны:

- проанализировать и запомнить основные определения и факты;
- повторить законспектированный на лекционном занятии материал и дополнить его с учетом рекомендованной по данной теме литературы;
- изучить рекомендованную основную и дополнительную литературу, составить тезисы, аннотации и конспекты наиболее важных моментов;
- самостоятельно выполнить домашние задания, аналогичные предлагаемым на занятиях;
- подготовить реферат по одной из рекомендуемых тем;
- подготовить ответы на вопросы к зачету.

8.1. Выполнение домашнего задания

Домашнее задание оценивается по следующим критериям:

- степень и уровень выполнения задания;
- аккуратность в оформлении работы;
- использование специальной литературы;
- сдача домашнего задания в срок.

8.2 Работа над рефератом

Реферат – индивидуальная письменная работа обучающегося, предполагающая анализ изложения в научных и других источниках определенной научной проблемы или вопроса.

Написание реферата практикуется в учебном процессе в целях приобретения обучающимися необходимой профессиональной подготовки, развития умения и навыков самостоятельного научного поиска: изучения литературы по выбранной теме, анализа различных источников и точек зрения, обобщения материала, выстраивания логики изложения, выделения главного, формулирования выводов.

Содержание реферата обучающийся представляет на консультациях. Предварительно подготовив тезисы доклада, обучающийся в течение 10-15 минут должен кратко изложить основные положения своей работы. После доклада автор отвечает на

вопросы, затем выступают оппоненты, которые заранее познакомились с текстом реферата, и отмечают его сильные и слабые стороны.

Как правило, реферат имеет стандартную структуру: титульный лист, содержание, введение, основное содержание темы, заключение, список использованных источников, приложения.

Оценивается оригинальность реферата, актуальность и полнота использованных источников, системность излагаемого материала, логика изложения и убедительность аргументации, оформление, своевременность срока сдачи, защита реферата перед аудиторией.

8.3. Подготовка к зачету

Подготовка к зачету предполагает:

- изучение основной и дополнительной литературы;
- изучение конспектов лекций;
- изучение задач, решенных на лабораторных работах и дома.

9. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине

Применение цифровых технологий в рамках преподавания дисциплины предоставляет расширенные возможности по организации учебных занятий в условиях цифровизации образования и позволяет сформировать у обучающихся навыки применения цифровых сервисов и инструментов в повседневной жизни и профессиональной деятельности.

Для реализации этой цели в рамках изучения дисциплины могут применяться следующие цифровые инструменты и сервисы:

- для коммуникации с обучающимися: VK Мессенджер (https://vk.me/app?mt_click_id=mt-v7eix5-1660908314-1651141140) – мессенджер, распространяется по лицензии FreeWare; Яндекс.Телемост (<https://telemost.yandex.ru/>) – инструмент для организации и проведения видеовстреч, распространяется по лицензии Shareware; Сервис WEEEEK (<https://weeek.net/ru>) – сервис для коммуникации, распространяется по лицензии trialware; Сферум (<https://sferum.ru/?p=start>) – мессенджер, распространяется по лицензии FreeWare; ouGile (<https://ru.yougile.com/>) – система управления проектами и общения, планировщик задач, распространяется по лицензии trialware;

– для организации удаленной связи и видеоконференций: Webinar (<https://webinar.ru/>) – платформа для вебинаров, обучения, распространяется по лицензии trialware; Видеозвонки Mail.ru (<https://calls.mail.ru/>) – сервис для видеозвонков, распространяется по лицензии ShareWare; Яндекс.Телемост (<https://telemost.yandex.ru/>) – сервис для видеозвонков, распространяется по лицензии ShareWare;

- для планирования аудиторных и внеаудиторных мероприятий: Яндекс.Календарь (<https://calendar.yandex.ru/>) – онлайн календарь-планер, распространяется по лицензии ShareWare; Shtab (<https://shtab.app/>) – планировщик задач, распространяется по лицензии FreeWare; Сервис WEEEEK (<https://weeek.net/ru>), распространяется по лицензии trialware;

– для управления удаленной работой, командой: Сервис WEEEEK (<https://weeek.net/ru>) – сервис для управления командой, распространяется по лицензии trialware; VK WorkSpace (<https://biz.mail.ru/>) – платформа для совместной удаленной работы (почта, сервис для коммуникаций, хранилище), распространяется по лицензии trialware;

- для совместного использования файлов: Яндекс.Документы (<https://docs.yandex.ru/>) – инструмент для создания и совместного использования документов, распространяется по лицензии Shareware; Yandex Forms (<https://cloud.yandex.ru/services/forms>) – бесплатный сервис для создания форм для опроса, регистрации и т.д., распространяется по лицензии Shareware; @Облако (<https://cloud.mail.ru/>) – сервис для создания, хранения и совместного

использования файлов, распространяется по лицензии Shareware; Яндекс.Диск – сервис для хранения и совместного использования документов, распространяется по лицензии Shareware.

Для успешного овладения дисциплиной используются следующие информационные технологии обучения: при проведении практических занятий используются презентации материала в программе Microsoft Office (PowerPoint), выход на профессиональные сайты, использование аудиоматериалов и видеоматериалов различных интернет-ресурсов.

Для дистанционной поддержки дисциплины используется система управления образовательным контентом Moodle. Для работы в данной системе все обучающиеся на первом курсе получают индивидуальные логин и пароль для входа в систему, в которой размещаются : программа дисциплины, материалы для лекционных и иных видов занятий , задания, контрольные вопросы.

Университет обеспечен необходимым комплектом лицензионного программного обеспечения:

- операционная система Windows 7, License 49013351 УГЛУТУ Russia 2011-09-06, OPEN 68975925ZZE1309. Срок действия - бессрочно;

- операционная система Astra Linux Special Edition. Договор №Pr000013979/0385/22-ЕП-223-06 от 01.07.2022. Срок действия: бессрочно;

- операционная система Windows Server. Контракт на услуги по предоставлению лицензий на право использовать компьютерное обеспечение № 067/ЭА от 07.12.2020 года. Срок бессрочно;

- пакет прикладных программ Office Professional Plus 2010, License 49013351 УГЛУТУ Russia 2011-09-06, OPEN 68975925ZZE1309. Срок действия – бессрочно;

- пакет прикладных программ Р7-Офис.Профессиональный. Договор №Pr000013979/0385/22-ЕП-223-06 от 01.07.2022. Срок: бессрочно;

- система управления обучением LMS Moodle – программное обеспечение с открытым кодом, распространяется по лицензии GNU Public License (rus);

- браузер Yandex (<https://yandex.ru/promo/browser/>) – программное обеспечение распространяется по простой (неисключительной) лицензии;

- платформа для анализа данных Deductor (Loginom Company), Academic, бесплатная версия для образования (<https://basegroup.ru/deductor/choice>);

- KNIME (Konstanz Information Miner) (<https://www.knime.com/downloads>) – бесплатная платформа для анализа данных, отчетности и интеграции с открытым исходным кодом, распространяется в соответствии со стандартной общественной лицензией GNU.

10. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине

Реализация учебного процесса осуществляется в специальных учебных аудиториях университета для проведения занятий лекционного типа, занятий семинарского и лабораторного типа, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации. Все аудитории укомплектованы специализированной мебелью и техническими средствами обучения, служащими для представления учебной информации большой аудитории. При необходимости обучающимся предлагаются наборы демонстрационного оборудования и учебно-наглядных пособий, обеспечивающие тематические иллюстрации.

Самостоятельная работа обучающихся выполняется в специализированной аудитории, которая оборудована учебной мебелью, компьютерной техникой с возможностью подключения к сети "Интернет" и обеспечением доступа в электронную информационно-образовательную среду УГЛУТУ.

Есть помещение для хранения и профилактического обслуживания учебного оборудования.

Требования к аудиториям

Наименование специальных помещений и помещений для самостоятельной работы	Оснащенность специальных помещений и помещений для самостоятельной работы
Помещение для лекционных занятий, групповых и индивидуальных консультаций, текущей и промежуточной аттестации.	Переносные: - демонстрационное мультимедийное оборудование (ноутбук, проектор); - комплект электронных учебно-наглядных материалов (презентаций) на флеш-носителях, обеспечивающих тематические иллюстрации. Столы и стулья. Экран.
Помещение для лабораторных занятий	Столы компьютерные, стулья. Персональные компьютеры. Выход в Интернет, электронную информационную образовательную среду университета. Доска аудиторная (интерактивная); демонстрационное мультимедийное оборудование: проектор
Помещения для самостоятельной работы	Столы компьютерные, стулья. Персональные компьютеры. Выход в Интернет, электронную информационную образовательную среду университета.
Помещение для хранения и профилактического обслуживания оборудования	Переносное демонстрационное оборудование (мультимедийные проекторы, экраны, ноутбуки). Расходные материалы для ремонта и обслуживания техники. Шкаф (стеллаж) для хранения экспонатов, таблиц, раздаточного материала. Места для хранения оборудования.