

**Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное учреждение высшего
образования
«ИРКУТСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ ТЕХНИЧЕСКИЙ
УНИВЕРСИТЕТ»**

Структурное подразделение «Брикс кафедры»

УТВЕРЖДЕНА:
на заседании кафедры
Протокол №15 от 18 марта 2025 г.

Рабочая программа дисциплины

«ИНТЕЛЛЕКТУАЛЬНЫЙ АНАЛИЗ ДАННЫХ / DATA MINING»

Направление: 09.03.01 Информатика и вычислительная техника

Искусственный интеллект и компьютерные науки /Artificial Intelligence and Computer
Science

Квалификация: Бакалавр

Форма обучения: очная

Документ подписан простой
электронной подписью
Составитель программы:
Локтионова Елена
Александровна
Дата подписания: 18.06.2025

Документ подписан простой
электронной подписью
Утвердил: Киреенко Анна
Павловна
Дата подписания: 19.06.2025

Документ подписан простой
электронной подписью
Согласовал: Афанасьев
Александр Диомидович
Дата подписания: 19.06.2025

Год набора – 2025

Иркутск, 2025 г.

1 Перечень планируемых результатов обучения по дисциплине, соотнесённых с планируемыми результатами освоения образовательной программы

1.1 Дисциплина «Интеллектуальный анализ данных / Data Mining» обеспечивает формирование следующих компетенций с учётом индикаторов их достижения

Код, наименование компетенции	Код индикатора компетенции
ПКС-2 Способен приобретать систематические знания в области искусственного интеллекта, понимая специфику научного знания, основанного на научно-обоснованных фактах	ПКС-2.5
ПКС-3 Способность разрабатывать программные продукты, используя современные технологии в области искусственного интеллекта	ПКС-3.11

1.2 В результате освоения дисциплины у обучающихся должны быть сформированы

Код индикатора	Содержание индикатора	Результат обучения
ПКС-3.11	Способен применить методы измерения и оценки характеристик программного обеспечения	Знать отличия Data Mining от классических статистических методов анализа и OLAP-систем, типы закономерностей и сферы применения data mining Уметь квалифицировать задачи data mining, применять методы интеллектуального анализа данных Владеть технологиями интеллектуального анализа данных, стандартами и инструментами, навыками анализа данных различной природы
ПКС-2.5	Владеет навыками приобретения знаний и умений в сфере интеллектуального анализа данных	Знать принципы и алгоритмы интеллектуального анализа данных для оптимизации документопотоков Уметь пользоваться современными программными инструментами для анализа эффективности систем организации хранения документов Владеть методами оптимизации документопотоков с использованием интеллектуальных систем и алгоритмов

2 Место дисциплины в структуре ООП

Изучение дисциплины «Интеллектуальный анализ данных / Data Mining» базируется на результатах освоения следующих дисциплин/практик: «Программирование / Programming», «Большие данные, методы обработки / Big Data, Processing Methods», «Основы проектной деятельности / Introduction to Project Development»

Дисциплина является предшествующей для дисциплин/практик: «Коммерциализация результатов интеллектуальной деятельности / Commercialization of Intellectual Property», «Производственная практика: преддипломная практика / Manufacturing Practice: Undergraduate Practice»

3 Объем дисциплины

Объем дисциплины составляет – 5 ЗЕТ

Вид учебной работы	Трудоемкость в академических часах (Один академический час соответствует 45 минутам астрономического часа)	
	Всего	Семестр № 7
Общая трудоемкость дисциплины	180	180
Аудиторные занятия, в том числе:	60	60
лекции	30	30
лабораторные работы	0	0
практические/семинарские занятия	30	30
Самостоятельная работа (в т.ч. курсовое проектирование)	84	84
Трудоемкость промежуточной аттестации	36	36
Вид промежуточной аттестации (итогового контроля по дисциплине)	Экзамен	Экзамен

4 Структура и содержание дисциплины

4.1 Сводные данные по содержанию дисциплины

Семестр № 7

№ п/п	Наименование раздела и темы дисциплины	Виды контактной работы						СРС		Форма текущего контроля
		Лекции		ЛР		ПЗ(СЕМ)		№	Кол. Час.	
		№	Кол. Час.	№	Кол. Час.	№	Кол. Час.			
1	2	3	4	5	6	7	8	9	10	11
1	Обзор задач интеллектуального анализа данных	1	4	1, 2, 3, 4, 5	30	3	6	2	18	Устный опрос
2	Основные методы интеллектуального анализа данных	2	8			1	6	1	16	Отчет
3	Процесс Data Mining	3	4			4	6	1	16	Отчет
4	Инструментальные средства интеллектуального анализа данных	4	6			5	6	1	16	Устный опрос
5	Практика применения интеллектуальных технологий	5	8			2	6	3	18	Отчет
	Промежуточная аттестация								36	Экзамен
	Всего		30		30		30		120	

4.2 Краткое содержание разделов и тем занятий

Семестр № 7

№	Тема	Краткое содержание
1	Обзор задач интеллектуального анализа данных	Понятие Интеллектуального анализа данных (Data Mining). Data Mining как часть рынка интеллектуальных технологий. Набор данных и их атрибутов. Измерения. Типы наборов данных. Форматы хранения данных. Метаданные.
2	Основные методы интеллектуального анализа данных	Задача классификации. Процесс классификации. Методы, применяемые для решения задач классификации. Точность классификации: оценка уровня ошибок. Оценивание классификационных методов. Деревья решений. Процесс конструирования дерева решений. Метод опорных векторов. Метод «ближайшего соседа». Байесова классификация. Задача прогнозирования. Сравнение задач прогнозирования и классификации. Прогнозирование и временные ряды. Решение задачи прогнозирования. Задача кластеризации. Применение кластерного анализа. Иерархичные методы. Итеративные методы. Методы поиска ассоциативных правил. Методы визуализации. Качество визуализации. Представление пространственных характеристик. Основные тенденции в визуализации.
3	Процесс Data Mining	Средства извлечения данных: методы и возможности. Начальные этапы: анализ предметной области; постановка задачи, подготовка данных. Очистка данных. Инструменты очистки данных. Построение и использование модели. Стандарты Data Mining.
4	Инструментальные средства интеллектуального анализа данных	Рынок инструментов Data Mining. Классификация инструментов Data Mining. Программное обеспечение для решения задач классификации. Программное обеспечения для решения задач кластеризации и сегментации. Программное обеспечение Data Mining для поиска ассоциативных правил. Программное обеспечение для решения задач оценивания и прогнозирования.
5	Практика применения интеллектуальных технологий	Системы бизнес-интеллекта и управления знаниями. Сферы применения Data Mining. Применение Data Mining для бизнес-задач. Data Mining для научных исследований. Data Mining консалтинг. Data Mining услуги. Примеры решения. Техническое описание решения. Технологии лингвистического анализа бизнес-информации. Интеллектуальный поиск в интернете. Аналитическая обработка бизнес-информации. Комплексный подход к внедрению Data Mining, OLAP и хранилищ данных.

		Интеграция OLAP и Data Mining. Хранилища данных. Преимущества хранилища данных.
--	--	---

4.3 Перечень лабораторных работ

Семестр № 7

№	Наименование лабораторной работы	Кол-во академических часов
1	Чтение и парсинг структурированных и полуструктурированных данных с помощью регулярных выражений	6
2	Применение методов классификации, кластерного анализа предобработанных данных и оценка уровня их ошибок	6
3	Построение и использование модели анализа данных, визуализация характеристик модели	6
4	Решение задач поиска ассоциативных правил и оценивания и прогнозирования на выбранном инструментальном ПО	6
5	Построение системы лингвистического анализа бизнес-информации	6

4.4 Перечень практических занятий

Семестр № 7

№	Темы практических (семинарских) занятий	Кол-во академических часов
1	Разбор лекционного материала. Текущий контроль	6
2	Разбор лекционного материала. Текущий контроль	6
3	Разбор лекционного материала. Текущий контроль	6
4	Разбор лекционного материала. Текущий контроль	6
5	Разбор лекционного материала. Текущий контроль	6

4.5 Самостоятельная работа

Семестр № 7

№	Вид СРС	Кол-во академических часов
1	Выполнение компьютерных экспериментов и компьютерных лабораторных работ в дистанционном режиме	48
2	Выполнение письменных творческих работ (писем, докладов, сообщений, ЭССЕ)	18
3	Оформление отчетов по лабораторным и	18

В ходе проведения занятий по дисциплине используются следующие интерактивные методы обучения: В ходе проведения занятий по дисциплине используются следующие интерактивные методы обучения: Дискуссии по проблемным вопросам, деловая игра, метод кейсов, метод проектов, круглый стол, просмотр и обсуждение учебных фильмов, блиц-опросы. Отдельные занятия по курсу могут проводиться в форме активного практического обучения: выездных занятий с посещением организаций и мероприятий для получения новых знаний и/или повторения материала на практике. При проведении таких занятий преподаватель выступает в качестве помощника и координатора процесса, передавая активную функцию обучения студентам. Он же регулирует процесс посредством подготовки специальных заданий, проведения консультаций, оценки знаний, умений и навыков, предоставления обратной связи. Помимо получения знаний активные практические занятия развивают коммуникативные навыки, учат студентов работать в команде, решать проблемы

5 Перечень учебно-методического обеспечения дисциплины

5.1 Методические указания для обучающихся по освоению дисциплины

5.1.1 Методические указания для обучающихся по практическим занятиям

Изучать дисциплину рекомендуется в соответствии с той последовательностью, которая обозначена в ее содержании. На лекциях преподаватель озвучивает тему, знакомит с перечнем литературы по теме, обосновывает место и роль этой темы в данной дисциплине, раскрывает ее практическое значение.

Практические (семинарские) занятия по своему содержанию связаны с тематикой лекционных занятий. Начинать подготовку к занятию целесообразно с конспекта лекций. Задание на практическое (семинарское) занятие сообщается обучающимся до его проведения. На семинаре преподаватель организует обсуждение этой темы, выступая в качестве организатора, консультанта и эксперта учебно-познавательной деятельности обучающегося.

5.1.2 Методические указания для обучающихся по самостоятельной работе:

Основными видами самостоятельной работы студентов с участием преподавателей являются:

- текущие консультации;
- выполнение письменных творческих работ (писем, докладов, сообщений, ЭССЕ);

Основными видами самостоятельной работы студентов без участия преподавателей являются:

- формирование и усвоение содержания конспекта лекций на базе рекомендованной лектором учебной литературы, включая информационные образовательные ресурсы (электронные учебники, электронные библиотеки и др.);
- самостоятельное изучение отдельных тем или вопросов по учебникам или учебным пособиям;
- подготовка к семинарам.

6 Фонд оценочных средств для контроля текущей успеваемости и проведения промежуточной аттестации по дисциплине

6.1 Оценочные средства для проведения текущего контроля

6.1.1 семестр 7 | Устный опрос

Описание процедуры.

Устный опрос предполагает неподготовленный ответ на одну из тем изучаемого раздела и ответы на уточняющие вопросы преподавателя. Ответ студента должен представлять собой логически последовательное сообщение на заданную тему, показывать его умение применять определения, правила в конкретных случаях.

Пример задания для устного опроса по теме "Обзор задач интеллектуального анализа данных":

1. Какие основные задачи решает интеллектуальный анализ данных (Data Mining) и в чем их отличие?
2. Что такое кластеризация и в каких ситуациях она применяется?
3. Каковы основные этапы процесса интеллектуального анализа данных и почему каждый из них важен?

Критерии оценивания.

Оценка «отлично» ставится при исчерпывающем ответе не содержащем ошибок, включающем привлечение материала из дополнительной литературы, который не входит в состав лекционного материала.

Оценка «хорошо» ставится при полном ответе, содержащем 1-2 несущественных ошибки, но не включающем материал из дополнительной литературы,

Оценка «удовлетворительно» ставится минимальном ответе, раскрывающем суть вопроса, содержащем 3-5 несущественных ошибок

6.1.2 семестр 7 | Отчет

Описание процедуры.

Подготовка и защита отчета о решении поставленной задачи.

Пример задачи для подготовки отчета по теме "Процесс Data Mining".

«Расписание курсов китайского языка»

Введение: Три преподавателя-волонтера из Китая будут преподавать китайский язык студентам ИРНИТУ. По предварительным опросам, более 200 студентов ИРНИТУ хотят изучать китайский язык. Для составления расписания курсов была собрана следующая информация о студентах: имя, номер группы, уровень образования, уровень китайского языка. Эта информация представлена в документе.

Задание: Составить расписание курсов китайского языка для студентов из списка, принимая во внимание следующие ограничения:

1. Курсы будут вести три преподавателя.
2. Студенты должны быть разделены на десять групп в соответствии с уровнем китайского языка.
3. Курсы могут проводиться с понедельника по субботу с 8.15 до 9.45, с 17.10 до 18.40 и с 18.45 до 20.15.
4. Занятия для магистрантов могут проводиться только в вечернее время.
5. Каждая группа студентов должна иметь 2 занятия в неделю.
6. Студенты ИРНИТУ должны иметь возможность посещать курсы китайского языка и не должны пропускать занятия по основному расписанию.

Дополнительную информацию (например, об основном расписании для студентов) можно найти на официальном сайте ИРНИТУ <https://www.istu.edu/schedule/>

Отчет о выполнении задания включает в себя следующие документы:

1. Документ, описывающий методику анализа имеющейся информации.
2. Пошаговое описание ваших действий и полученных промежуточных результатов.
3. Список студенческих групп, соответствующих вышеуказанным критериям
4. Расписание курсов китайского языка, соответствующих вышеуказанным критериям

Критерии оценивания.

Отлично: Представлен полный комплект необходимых документов в соответствии с заданием, документы заполнены правильно и содержание документов соответствует предъявляемым требованиям. Выборочная проверка не показывает наложений расписания курсов китайского языка на основное расписание

Хорошо: Представлен полный комплект необходимых документов в соответствии с заданием, документы заполнены правильно и содержание документов.

соответствует предъявляемым требованиям. Выборочная проверка показывает наличие незначительного числа наложений расписания курсов китайского языка на основное расписание.

Удовлетворительно: Представлен полный комплект необходимых документов в соответствии с заданием, документы заполнены с небольшим количеством ошибок и содержание документов в целом соответствует предъявляемым требованиям. Выборочная проверка показывает наличие значительного числа наложений расписания курсов китайского языка на основное расписание.

Неудовлетворительно: Представлен неполный комплект необходимых документов в соответствии с заданием, документы заполнены с ошибками, их содержание не соответствует предъявляемым требованиям. Выборочная проверка показывает наличие значительного числа наложений расписания курсов китайского языка на основное расписание.

6.2 Оценочные средства для проведения промежуточной аттестации

6.2.1 Критерии и средства (методы) оценивания индикаторов достижения компетенции в рамках промежуточной аттестации

Индикатор достижения компетенции	Критерии оценивания	Средства (методы) оценивания промежуточной аттестации
ПКС-3.11	Обнаруживает всестороннее, систематическое и глубокое знание учебно-программного материала; принимает активное участие в обсуждении по теме занятия; усвоил основную и дополнительную литературу, рекомендованную рабочей программой дисциплины; проявляет творческие способности в понимании, изложении и использовании учебно-программного материала.	Отчеты по лабораторным работам, вопросы для контроля по лабораторным работам, устный опрос
ПКС-2.5	Обнаруживает всестороннее, систематическое и глубокое знание учебно-программного материала;	Отчеты по лабораторным работам, вопросы

	принимает активное участие в обсуждении по теме занятия; усвоил основную и дополнительную литературу, рекомендованную рабочей программой дисциплины; проявляет творческие способности в понимании, изложении и использовании учебно-программного материала.	для контроля по лабораторным работам, устный опрос
--	---	--

6.2.2 Типовые оценочные средства промежуточной аттестации

6.2.2.1 Семестр 7, Типовые оценочные средства для проведения экзамена по дисциплине

6.2.2.1.1 Описание процедуры

Подготовка отчета (комплекта документов) о решении поставленной задачи и его публичная защита.

6.2.2.1.2 Критерии оценивания

Отлично	Хорошо	Удовлетворительно	Неудовлетворительно
Представлен полный комплект необходимых документов в соответствии с заданием, документы заполнены правильно и содержание документов соответствует предъявляемым требованиям. Во время защиты отчета студент полно излагает материал (отвечает на вопросы), дает правильное определение основных понятий; обнаруживает понимание	Представлен полный комплект необходимых документов в соответствии с заданием, документы заполнены правильно и содержание документов соответствует предъявляемым требованиям. Во время защиты студент дает ответ, удовлетворяющий тем же требованиям, что и для оценки «отлично», но допускает 1-2 ошибки, которые сам же исправляет.	Представлен полный комплект необходимых документов в соответствии с заданием, документы заполнены с небольшим количеством ошибок и содержание документов в целом соответствует предъявляемым требованиям. Во время защиты отчета студент обнаруживает знание и понимание основных положений данной темы, но излагает материал неполно и допускает неточности в определении	Представлен неполный комплект необходимых документов в соответствии с заданием, документы заполнены с ошибками, их содержание не соответствует предъявляемым требованиям. Во время защиты отчета студент обнаруживает незнание большей части соответствующего вопроса.

материала, может обосновать свои суждения, применить знания на практике.		понятий или формулировке правил; не умеет достаточно глубоко и доказательно обосновать свои суждения и привести свои примеры.	
--	--	---	--

7 Основная учебная литература

1. Сапрыкин О. Н. Интеллектуальный анализ данных : учебное пособие / О. Н. Сапрыкин, 2020. - 80.
2. Salcedo J. Machine Learning for Data Mining / J. Salcedo, 2019. - 247.
3. Integration Challenges for Analytics, Business Intelligence, and Data Mining / ed.: A. Azevedo, M. F. Santos, 2020. - 271.
4. Ghavami P. Big Data Analytics Methods. Analytics Techniques in Data Mining, Deep Learning and Natural Language Processing / P. Ghavami, 2019. - 255.

8 Дополнительная учебная литература и справочная

1. Нестеров С. А. Интеллектуальный анализ данных с использованием SQL Server : учебник для вузов / С. А. Нестеров, 2024. - 160.

9 Ресурсы сети Интернет

1. <http://library.istu.edu/>
2. <https://e.lanbook.com/>

10 Профессиональные базы данных

1. <http://new.fips.ru/>
2. <http://www1.fips.ru/>

11 Перечень информационных технологий, лицензионных и свободно распространяемых специализированных программных средств, информационных справочных систем

12 Материально-техническое обеспечение дисциплины

1. Компьютер Intel Core i3 3220/DDR3 4Gb/Hdd 1Tb/GF 1Gb/DVD+RW/Sound Net km/LCD23"/