

МИНОБРНАУКИ РОССИИ
ВЛАДИВОСТОКСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
НАУЧНО-ОБРАЗОВАТЕЛЬНЫЙ ЦЕНТР "ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ"

Рабочая программа дисциплины (модуля)
ТЕХНОЛОГИИ СБОРА И ПРЕДВАРИТЕЛЬНОЙ ОБРАБОТКИ ДАННЫХ

Направление и направленность (профиль)
09.04.03 Прикладная информатика. Интеллектуальный анализ данных

Год набора на ОПОП
2025

Форма обучения
очная

Владивосток 2025

Рабочая программа дисциплины (модуля) «Технологии сбора и предварительной обработки данных» составлена в соответствии с требованиями ФГОС ВО по направлению подготовки 09.04.03 Прикладная информатика (утв. приказом Минобрнауки России от 19.09.2017г. №916) и Порядком организации и осуществления образовательной деятельности по образовательным программам высшего образования – программам бакалавриата, программам специалитета, программам магистратуры (утв. приказом Минобрнауки России от 06.04.2021 г. N245).

Составитель(и):

Назаров Д.А., кандидат технических наук, доцент, Научно-образовательный центр "Искусственный интеллект"

Утверждена на заседании научно-образовательный центр "искусственный интеллект" от 05.06.2025 , протокол № 6

СОГЛАСОВАНО:

Заведующий кафедрой (разработчика)

Кригер А.Б.

ДОКУМЕНТ ПОДПИСАН ЭЛЕКТРОННОЙ ПОДПИСЬЮ	
Сертификат	1582918206
Номер транзакции	0000000000DF292C
Владелец	Кригер А.Б.

1 Цель, планируемые результаты обучения по дисциплине (модулю)

Целью освоения дисциплины являются формирование у студентов магистратуры компетенции в области автоматизации сбора и обработки информации.

Основные задачи изучения дисциплины:

- формирование углублённых знаний об основных методах использования современных информационных технологий для сбора и обработки данных;
- формирование практических навыков:

1.

1. создания автоматизированных информационных систем по сбору и обработке информации;
2. поиска, анализа и оценки источников информации для обеспечения возможности их сбора в автоматическом режиме;
3. выбор наиболее подходящих инструментов для создания информационных систем по сбору и обработке информации.

Планируемыми результатами обучения по дисциплине (модулю), являются знания, умения, навыки. Перечень планируемых результатов обучения по дисциплине (модулю), соотнесенных с планируемыми результатами освоения образовательной программы, представлен в таблице 1.

Таблица 1 – Компетенции, формируемые в результате изучения дисциплины (модуля)

Название ОПОП ВО, сокращенное	Код и формулировка компетенции	Код и формулировка индикатора достижения компетенции	Результаты обучения по дисциплине		
			Код результата	Формулировка результата	
09.04.03 «Прикладная информатика» (М-ПИ)	ОПК-6 : Способен исследовать современные проблемы и методы прикладной информатики и развития информационного общества	ОПК-6.1к : Выявляет технологии в сфере прикладной информатики, влияющие на развитие информационного общества	РД1	Знание	методологию системного подхода
			РД8	Умение	применять на практике перспективных методик исследования прикладных и информационных процессов
			РД9	Навык	навыками логико-методологического анализа научного исследования
		ОПК-6.2к : Разрабатывает методы оценки качества информации и технологии представления информации и данных	РД10	Знание	теоретических проблем прикладной информатики, в том числе семантической обработки информации, развития представлений об оценке качества информации в информационных системах
			РД11	Умение	проводить анализ структуры информационных ресурсов
			РД6	Навык	методами обеспечения качества и развития процесса разработки программ

		ОПК-6.3к : Решает задачи различных классов, используя современные методы, средства и стандарты прикладной информатики, и проводит их анализ	РД10	Знание	теоретических проблем прикладной информатики, в том числе семантической обработки информации, развития представлений об оценке качества информации в информационных системах
			РД8	Умение	применять на практике перспективных методик исследования прикладных и информационных процессов
			РД9	Навык	навыками логико-методологического анализа научного исследования

В процессе освоения дисциплины решаются задачи воспитания гармонично развитой, патриотичной и социально ответственной личности на основе традиционных российских духовно-нравственных и культурно-исторических ценностей, представленные в таблице 1.2.

Таблица 1.2 – Целевые ориентиры воспитания

Воспитательные задачи	Формирование ценностей	Целевые ориентиры
Формирование научного мировоззрения и культуры мышления		
Формирование осознания ценности научного мировоззрения и критического мышления	Гуманизм	Системное мышление
Формирование коммуникативных навыков и культуры общения		
Формирование навыков публичного выступления и презентации своих идей	Взаимопомощь и взаимоуважение	Умение работать в команде и взаимопомощь

2 Место дисциплины (модуля) в структуре ОПОП

Дисциплина относится к базовой части Блока 1 Дисциплины (Модули)

3. Объем дисциплины (модуля)

Объем дисциплины (модуля) в зачетных единицах с указанием количества академических часов, выделенных на контактную работу с обучающимися (по видам учебных занятий) и на самостоятельную работу, приведен в таблице 2.

Таблица 2 – Общая трудоемкость дисциплины

Название ОПОП ВО	Форма обуче- ния	Часть УП	Семестр (ОФО) или курс (ЗФО, ОЗФО)	Трудо- емкость	Объем контактной работы (час)					СРС	Форма аттес- тации	
				(З.Е.)	Всего	Аудиторная			Внеауди- торная			
						лек.	прак.	лаб.	ПА			КСР

09.04.03 Прикладная информатика	ОФО	M01.Б	1	4	33	8	24	0	1	0	111	Э
---------------------------------------	-----	-------	---	---	----	---	----	---	---	---	-----	---

4 Структура и содержание дисциплины (модуля)

4.1 Структура дисциплины (модуля) для ОФО

Тематический план, отражающий содержание дисциплины (перечень разделов и тем), структурированное по видам учебных занятий с указанием их объемов в соответствии с учебным планом, приведен в таблице 3.1

Таблица 3.1 – Разделы дисциплины (модуля), виды учебной деятельности и формы текущего контроля для ОФО

№	Название темы	Код результата обучения	Кол-во часов, отведенное на				Форма текущего контроля
			Лек	Практ	Лаб	СРС	
1	Информация в современном мире	РД1, РД6, РД7	2	0	0	8	не предусмотрен
2	Источники и способы хранения цифровой информации	РД4, РД5, РД6, РД7	2	0	0	10	не предусмотрен
3	Методы и средства сбора данных	РД1, РД6, РД8, РД9	2	0	0	10	не предусмотрен
4	Методы анализа данных	РД2, РД3, РД7, РД8, РД9	2	0	0	10	не предусмотрен
5	Работа с пропущенными данными	РД3, РД7, РД8	0	4	0	12	проектное задание
6	Веб-технологии хранения, обработки и передачи информации	РД2, РД5, РД6, РД8	0	8	0	20	проектное задание
7	API-интерфейсы предоставления информации	РД4, РД6, РД8, РД9	0	6	0	20	проектное задание
8	Структурированные виды данных	РД5, РД7	0	6	0	21	проектное задание
Итого по таблице			8	24	0	111	

4.2 Содержание разделов и тем дисциплины (модуля) для ОФО

Тема 1 Информация в современном мире.

Содержание темы: 1.1. Основные определения информации. Типы и виды носителей и накопителей информации. 1.2. Жизненный цикл информации. Преобразования информации в цифровой вид.

Формы и методы проведения занятий по теме, применяемые образовательные технологии: Лекция.

Виды самостоятельной подготовки студентов по теме: теоретическая подготовка, компьютерное моделирование.

Тема 2 Источники и способы хранения цифровой информации.

Содержание темы: 2.1. Информационные ресурсы. 2.2. Базы данных. 2.3. Веб-ресурсы. 2.4. Распределенные информационные системы.

Формы и методы проведения занятий по теме, применяемые образовательные технологии: Лекция.

Виды самостоятельной подготовки студентов по теме: теоретическая подготовка, компьютерное моделирование.

Тема 3 Методы и средства сбора данных.

Содержание темы: 3.1. Особенности хранения информации на локальных и глобальных ресурсах. 3.2. Глобальные информационные ресурсы. 3.3. Веб-технологии хранения, обработки и передачи информации. 3.4. API-интерфейсы предоставления информации. 3.5. Структурированные виды данных.

Формы и методы проведения занятий по теме, применяемые образовательные технологии: Лекция.

Виды самостоятельной подготовки студентов по теме: теоретическая подготовка, компьютерное моделирование.

Тема 4 Методы анализа данных.

Содержание темы: 4.1. Распределения вероятностей. 4.2. Статистика выводов. 4.3. Анализ табличных данных. 4.4. Основы регрессионного анализа. 4.5. Интеллектуальный анализ данных.

Формы и методы проведения занятий по теме, применяемые образовательные технологии: Лекция.

Виды самостоятельной подготовки студентов по теме: теоретическая подготовка, компьютерное моделирование.

Тема 5 Работа с пропущенными данными.

Содержание темы: 1. Обнаружения пропущенных, недопустимых и бесконечных значений с помощью функции `is.na()`, `is.nan()` и `is.infinite()` программы R-Studio. 2. Изучение структуры пропусков в соотношении с другими переменными массива с помощью функции `md.pattern()` библиотеки `mice` программы R-Studio. 3. Функции для визуализации структуры пропущенных данных `aggr()`, `matrixplot()` и `scattMiss()` пакета `VIM` программы R-Studio. 4. Обнаружения строк в данных, которые не содержат пропущенных значений с помощью функции `complete.cases()`. Построчное удаление. 5. Простое восстановление данных с помощью пакета `Hmisc`. 6. Многомерное восстановление данных с помощью пакета `mice`. Моделирование набора массивов для пропущенных значений с помощью функции `mice()`. Применение необходимого статистического метода с помощью функции `with()` (например, линейная регрессия). Объединение результатов с помощью функции `pool()`. 7. Специализированные методы работы с пропущенными данными. Оценка методом максимального правдоподобия для многомерных нормальных данных с пропущенными значениями с помощью пакета `Mvnmle`.

Формы и методы проведения занятий по теме, применяемые образовательные технологии: Практическое занятие.

Виды самостоятельной подготовки студентов по теме: теоретическая подготовка, компьютерное моделирование.

Тема 6 Веб-технологии хранения, обработки и передачи информации.

Содержание темы: 1. Анализ структуры HTML документа. 2. Программная реализация GET и POST-запросов. 3. Разбор DOM-дерева HTML документа. 4. Автоматизация сбора данных из открытого веб-ресурса.

Формы и методы проведения занятий по теме, применяемые образовательные технологии: Практическое занятие.

Виды самостоятельной подготовки студентов по теме: теоретическая подготовка, компьютерное моделирование.

Тема 7 API-интерфейсы предоставления информации.

Содержание темы: 1. Структура и формат данных XML и JSON. 2. Генерация и разбор XML-документа. 3. Реализация сервера на основе RESTful API. 4. Реализация интерфейса взаимодействия с API-интерфейсом.

Формы и методы проведения занятий по теме, применяемые образовательные технологии: Практическое занятие.

Виды самостоятельной подготовки студентов по теме: теоретическая подготовка, компьютерное моделирование.

Тема 8 Структурированные виды данных.

Содержание темы: 1. Структура и формат документов в формате MS Office. 2. COM-объекты MS Office. 3. Автоматизация анализа файлов формата MS Word и Excel.

Формы и методы проведения занятий по теме, применяемые образовательные технологии: Практическое занятие.

Виды самостоятельной подготовки студентов по теме: теоретическая подготовка, компьютерное моделирование.

5 Методические указания для обучающихся по изучению и реализации дисциплины (модуля)

5.1 Методические рекомендации обучающимся по изучению дисциплины и по обеспечению самостоятельной работы

5.2 Особенности организации обучения для лиц с ограниченными возможностями здоровья и инвалидов

При необходимости обучающимся из числа лиц с ограниченными возможностями здоровья и инвалидов (по заявлению обучающегося) предоставляется учебная информация в доступных формах с учетом их индивидуальных психофизических особенностей:

- для лиц с нарушениями зрения: в печатной форме увеличенным шрифтом; в форме электронного документа; индивидуальные консультации с привлечением тифлосурдопереводчика; индивидуальные задания, консультации и др.

- для лиц с нарушениями слуха: в печатной форме; в форме электронного документа; индивидуальные консультации с привлечением сурдопереводчика; индивидуальные задания, консультации и др.

- для лиц с нарушениями опорно-двигательного аппарата: в печатной форме; в форме электронного документа; индивидуальные задания, консультации и др.

6 Фонд оценочных средств для проведения текущего контроля и промежуточной аттестации обучающихся по дисциплине (модулю)

В соответствии с требованиями ФГОС ВО для аттестации обучающихся на соответствие их персональных достижений планируемым результатам обучения по дисциплине (модулю) созданы фонды оценочных средств. Типовые контрольные задания, методические материалы, определяющие процедуры оценивания знаний, умений и навыков, а также критерии и показатели, необходимые для оценки знаний, умений, навыков и характеризующие этапы формирования компетенций в процессе освоения образовательной программы, представлены в Приложении 1.

7 Учебно-методическое и информационное обеспечение дисциплины (модуля)

7.1 Основная литература

1. Анализ данных : учебник для вузов / под редакцией В. С. Мхитаряна. — Москва : Издательство Юрайт, 2025. — 448 с. — (Высшее образование). — ISBN 978-5-534-19964-2. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/560311> (дата обращения: 18.06.2025).
2. Демидова, Л. А. Интеллектуальный анализ данных на языке Python : учебно-методическое пособие / Л. А. Демидова. — Москва : РТУ МИРЭА, 2021. — 92 с. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/218693> (дата обращения: 17.06.2025). — Режим доступа: для авториз. пользователей.
3. Криволапов, С. Я. Введение в анализ данных. Поиск структуры данных с применением языка Python : учебное пособие / С.Я. Криволапов. — Москва : ИНФРА-М, 2024. — 177 с. — (Высшее образование). — DOI 10.12737/2082643. - ISBN 978-5-16-019001-3. - Текст : электронный. - URL: <https://znanium.ru/catalog/product/2082643> (Дата обращения -18.06.2025)

7.2 Дополнительная литература

1. Минязев, Р. Ш. Скриптовые языки web-программирования (JavaScript, PHP, html/CSS) : учебно-методическое пособие / Р. Ш. Минязев. — Казань : КНИТУ-КАИ, 2022. — 60 с. — ISBN 978-5-7579-2632-2. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/399557> (дата обращения: 17.06.2025). — Режим доступа: для авториз. пользователей.
2. Нурмагомедова, Н. Х. Лабораторный практикум по дисциплине «Web-технологии» : учебное пособие / Н. Х. Нурмагомедова. — Махачкала : ДГПУ, 2023 — Часть 1 : Язык HTML — 2023. — 53 с. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/406952> (дата обращения: 17.06.2025). — Режим доступа: для авториз. пользователей.
3. Соловьева, О. М. Web-программирование : учебно-методическое пособие / О. М. Соловьева. — Санкт-Петербург : СПбГУТ им. М.А. Бонч-Бруевича, 2023. — 123 с. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/425960> (дата обращения: 17.06.2025). — Режим доступа: для авториз. пользователей.
4. Тузовский, А. Ф. Проектирование и разработка web-приложений : учебник для вузов / А. Ф. Тузовский. — Москва : Издательство Юрайт, 2025. — 219 с. — (Высшее образование). — ISBN 978-5-534-16300-1. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/561176> (дата обращения: 18.06.2025).

7.3 Ресурсы информационно-телекоммуникационной сети "Интернет", включая профессиональные базы данных и информационно-справочные системы (при необходимости):

1. Образовательная платформа "ЮРАЙТ"
2. Электронно-библиотечная система "ZNANIUM.COM"
3. Электронно-библиотечная система "ЛАНЬ"
4. Open Academic Journals Index (ОАИ). Профессиональная база данных - Режим доступа: <http://oaji.net/>
5. Президентская библиотека им. Б.Н.Ельцина (база данных различных профессиональных областей) - Режим доступа: <https://www.prilib.ru/>

6. Информационно-справочная система "Консультант Плюс" - Режим доступа:
<http://www.consultant.ru/>

8 Материально-техническое обеспечение дисциплины (модуля) и перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине (модулю), включая перечень программного обеспечения

Основное оборудование:

- Коммутатор SuperStack 3 (16*10/100 19")
- Монитор облачный 23" LG23CAV42K/мышь Genius
- Мультимедийный проектор №1 Casio XJ-V2
- Облачный монитор 23" LG CAV42K
- Облачный монитор LG Electronics черный +клавиатура+мышь
- П/К DNS Office T300, мышь Genius NetScroll 100, клавиатура Genius KB-06X, монитор AOC919 19"
- Проектор Casio XJ-V1
- Уст-во бесп.питания UPS-3000

Программное обеспечение:

- □ Microsoft Office 2010 Standart
- □ Microsoft Windows 7 Russian
- □ RStudio
- □ Visual Studio
- □ СПС КонсультантПлюс: Версия Проф

МИНОБРНАУКИ РОССИИ

ВЛАДИВОСТОКСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

НАУЧНО-ОБРАЗОВАТЕЛЬНЫЙ ЦЕНТР "ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ"

Фонд оценочных средств
для проведения текущего контроля
и промежуточной аттестации по дисциплине (модулю)

ТЕХНОЛОГИИ СБОРА И ПРЕДВАРИТЕЛЬНОЙ ОБРАБОТКИ ДАННЫХ

Направление и направленность (профиль)
09.04.03 Прикладная информатика. Интеллектуальный анализ данных

Год набора на ОПОП
2025

Форма обучения
очная

Владивосток 2025

1 Перечень формируемых компетенций

Название ОПОП ВО, сокращенное	Код и формулировка компетенции и	Код и формулировка индикатора достижения компетенции
09.04.03 «Прикладная информатика» (М-ПИ)	ОПК-6 : Способен исследовать современные проблемы и методы прикладной информатики и развития информационного общества	ОПК-6.1к : Выявляет технологии в сфере прикладной информатики, влияющие на развитие информационного общества
		ОПК-6.2к : Разрабатывает методы оценки качества информации и технологии представления информации и данных
		ОПК-6.3к : Решает задачи различных классов, используя современные методы, средства и стандарты прикладной информатики, и проводит их анализ

Компетенция считается сформированной на данном этапе в случае, если полученные результаты обучения по дисциплине оценены положительно (диапазон критериев оценивания результатов обучения «зачтено», «удовлетворительно», «хорошо», «отлично»). В случае отсутствия положительной оценки компетенция на данном этапе считается несформированной.

2 Показатели оценивания планируемых результатов обучения

Компетенция ОПК-6 «Способен исследовать современные проблемы и методы прикладной информатики и развития информационного общества»

Таблица 2.1 – Критерии оценки индикаторов достижения компетенции

Код и формулировка индикатора достижения компетенции	Результаты обучения по дисциплине			Критерии оценивания результатов обучения
	Код	Тип	Результат	
ОПК-6.1к : Выявляет технологии в сфере прикладной информатики, влияющие на развитие информационного общества	РД 1	Знание	методологию системного подхода	знает методологию системного подхода
	РД 8	Умение	применять на практике перспективных методик исследования прикладных и информационных процессов	умеет применять на практике перспективных методик исследования прикладных и информационных процессов
	РД 9	Навык	навыками логико-методологического анализа научного исследования	обладает навыками логико-методологического анализа научного исследования
ОПК-6.2к : Разрабатывает методы оценки качества информации и технологии представления информации и данных	РД 6	Навык	методами обеспечения качества и развития процесса разработки программ	обладание методами обеспечения качества и развития процесса разработки программ
	РД 10	Знание	теоретических проблем прикладной информатики, в том числе семантической обработки информации, развития представлений об оценке качества информации в информационных системах	знает теоретических проблем прикладной информатики, в том числе семантической обработки информации, развития представлений об оценке качества информации в информационных системах

	РД 11	Умение	проводить анализ структуры информационных ресурсов	умеет проводить анализ структуры информационных ресурсов
ОПК-6.3к : Решает задачи различных классов, используя со временные методы, средства и стандарты прикладной информатики, и проводит их анализ	РД 8	Умение	применять на практике перспективных методик исследования прикладных и информационных процессов	умеет применять на практике перспективных методик исследования прикладных и информационных процессов
	РД 9	Навык	навыками логико-методологического анализа научного исследования	обладает навыками логико-методологического анализа научного исследования
	РД 10	Знание	теоретических проблем прикладной информатики, в том числе семантической обработки информации, развития представлений об оценке качества информации в информационных системах	знает теоретических проблем прикладной информатики, в том числе семантической обработки информации, развития представлений об оценке качества информации в информационных системах

Таблица заполняется в соответствии с разделом 1 Рабочей программы дисциплины (модуля).

3 Перечень оценочных средств

Таблица 3 – Перечень оценочных средств по дисциплине (модулю)

Контролируемые планируемые результаты обучения		Контролируемые темы дисциплины	Наименование оценочного средства и представление его в ФОС	
			Текущий контроль	Промежуточная аттестация
Очная форма обучения				
РД1	Знание : методологию системного подхода	1.1. Информация в современном мире	не предусмотрен	Проект
		1.3. Методы и средства сбора данных	не предусмотрен	Проект
РД2	Умение : решать задачи, требующие навыков абстрактного мышления	1.4. Методы анализа данных	не предусмотрен	Проект
		1.6. Веб-технологии хранения, обработки и передачи информации	не предусмотрен	Проект
РД3	Навык : владение методами анализа и синтеза	1.4. Методы анализа данных	Проект	Проект
		1.5. Работа с пропущенными данными	Проект	Проект
РД4	Знание : концептуальные основы методологии функциональной декомпозиции программных систем	1.2. Источники и способы хранения цифровой информации	не предусмотрен	Проект
		1.7. API-интерфейсы предоставления информации	не предусмотрен	Проект
РД5	Умение : выполнить анализ требований и создание алгоритмов функцио	1.2. Источники и способы хранения цифровой информации	Проект	Проект

	нирования программного обеспечения	1.6. Веб-технологии хранения, обработки и передачи информации	Проект	Проект
		1.8. Структурированные виды данных	Проект	Проект
РД6	Навык : методами обеспечения качества и развития процесса разработки программ	1.1. Информация в современном мире	Проект	Проект
		1.2. Источники и способы хранения цифровой информации	Проект	Проект
		1.3. Методы и средства сбора данных	Проект	Проект
		1.6. Веб-технологии хранения, обработки и передачи информации	Проект	Проект
		1.7. API-интерфейсы предоставления информации	Проект	Проект
РД7	Знание : основных понятия методов интеллектуальной обработки данных и организации баз знаний	1.1. Информация в современном мире	не предусмотрен	Проект
		1.2. Источники и способы хранения цифровой информации	не предусмотрен	Проект
		1.4. Методы анализа данных	не предусмотрен	Проект
		1.5. Работа с пропущенными данными	не предусмотрен	Проект
		1.8. Структурированные виды данных	не предусмотрен	Проект
РД8	Умение : применять на практике перспективных методик исследования прикладных и информационных процессов	1.3. Методы и средства сбора данных	Проект	Проект
		1.4. Методы анализа данных	Проект	Проект
		1.5. Работа с пропущенными данными	Проект	Проект
		1.6. Веб-технологии хранения, обработки и передачи информации	Проект	Проект
		1.7. API-интерфейсы предоставления информации	Проект	Проект
РД9	Навык : навыками логико-методологического анализа научного исследования	1.3. Методы и средства сбора данных	не предусмотрен	Проект
		1.4. Методы анализа данных	не предусмотрен	Проект
		1.7. API-интерфейсы предоставления информации	не предусмотрен	Проект

4 Описание процедуры оценивания

Качество сформированности компетенций на данном этапе оценивается по результатам текущих и промежуточных аттестаций при помощи количественной оценки,

выраженной в баллах. Максимальная сумма баллов по дисциплине (модулю) равна 100 баллам.

Промежуточная аттестация по дисциплине «Технологии сбора и предварительной обработки данных» включает в себя посещение лекционных занятий, выполнение практических заданий и контрольных работ. На лекционных занятиях дается основной систематизированный теоретический материал. На практических занятиях под руководством преподавателя студенты получают навыки работы с программой R-Studio, а также приобретают умения создания программных комплексов автоматизации сбора данных из открытых источников информации. Самостоятельная работа студентов заключается в изучении предлагаемой литературы для усвоения и углубления полученных аудиторных знаний. Контрольные работы позволяют оценить уровень усвоения обучающимися знаний, умений и навыков и выявить степень сформированности компетенций. Предусмотренная форма аттестации – дифференцированный зачет.

Объем и качество освоения обучающимися дисциплины, уровень сформированности дисциплинарных компетенций оцениваются по результатам текущих и промежуточной аттестаций количественной оценкой, выраженной в баллах, максимальная сумма баллов по дисциплине равна 100 баллам.

Вид учебной деятельности	Оценочное средство	
	Проект	Итого
Лекции	10	10
Практические занятия	40	40
СРС	50	50
Итого	100	100

Сумма баллов, набранных студентом по всем видам учебной деятельности в рамках дисциплины, переводится в оценку в соответствии с таблицей

Сумма баллов по дисциплине	Оценка по промежуточной аттестации	Характеристика уровня освоения дисциплины
от 91 до 100	«отлично»	Студент демонстрирует сформированность дисциплинарной компетенции на итоговом уровне, обнаруживает всестороннее, систематическое и глубокое знание учебного материала, усвоил основную литературу и знаком с дополнительной литературой, рекомендованной программой, умеет свободно выполнять практические задания, предусмотренные программой, свободно оперирует приобретенными знаниями, умениями, применяет их в ситуациях повышенной сложности.
от 76 до 90	«хорошо»	Студент демонстрирует сформированность дисциплинарной компетенции на среднем уровне: основные знания, умения освоены, но допускаются незначительные ошибки, неточности, затруднения при аналитических операциях, переносе знаний и умений на новые, нестандартные ситуации.
от 61 до 75	«удовлетворительно»	Студент демонстрирует сформированность дисциплинарных компетенций на базовом уровне: в ходе контрольных мероприятий допускаются значительные ошибки, проявляется отсутствие отдельных знаний, умений, навыков по дисциплинарной компетенции, студент испытывает значительные затруднения при оперировании знаниями и умениями при их переносе на новые ситуации.
от 41 до 60	«не зачтено»	Студент демонстрирует сформированность дисциплинарных компетенций на уровне ниже базового, проявляется недостаточность знаний, умений, навыков.
от 0 до 40	«не зачтено»	Дисциплинарные компетенции не сформированы. Проявляется полное или практически полное отсутствие знаний, умений, навыков.

Сумма баллов, набранных студентом по всем видам учебной деятельности в рамках дисциплины, переводится в оценку в соответствии с таблицей.

Сумма баллов по дисциплине	Оценка по промежуточной аттестации	Характеристика качества сформированности компетенции
от 91 до 100	«зачтено» / «отлично»	Студент демонстрирует сформированность дисциплинарных компетенций, обнаруживает всестороннее, систематическое и глубокое знание учебного материала, усвоил основную литературу и знаком с дополнительной литературой, рекомендованной программой, умеет свободно выполнять практические задания, предусмотренные программой, свободно оперирует приобретенными знаниями, умениями, применяет их в ситуациях повышенной сложности.
от 76 до 90	«зачтено» / «хорошо»	Студент демонстрирует сформированность дисциплинарных компетенций: основные знания, умения освоены, но допускаются незначительные ошибки, неточности, затруднения при аналитических операциях, переносе знаний и умений на новые, нестандартные ситуации.
от 61 до 75	«зачтено» / «удовлетворительно»	Студент демонстрирует сформированность дисциплинарных компетенций: в ходе контрольных мероприятий допускаются значительные ошибки, проявляется отсутствие отдельных знаний, умений, навыков по некоторым дисциплинарным компетенциям, студент испытывает значительные затруднения при оперировании знаниями и умениями при их переносе на новые ситуации.
от 41 до 60	«не зачтено» / «неудовлетворительно»	У студента не сформированы дисциплинарные компетенции, проявляется недостаточность знаний, умений, навыков.
от 0 до 40	«не зачтено» / «неудовлетворительно»	Дисциплинарные компетенции не сформированы. Проявляется полное или практически полное отсутствие знаний, умений, навыков.

5 Примерные оценочные средства

5.1 Проект

Примерный перечень технологий обработки данных, реализуемых в рамках проектного задания

1. Обнаружения пропущенных, недопустимых и бесконечных значений с помощью функции `is.na()`, `is.nan()` и `is.infinite()` программы R - Studio.
2. Изучение структуры пропусков в соотношении с другими переменными массива с помощью функции `md.pattern()` библиотеки `mice` программы R - Studio.
3. Функции для визуализации структуры пропущенных данных `aggr()`, `matrixplot()` и `scattMiss()` пакета `VIM` программы R - Studio.
4. Обнаружения строк в данных, которые не содержат пропущенных значений с помощью функции `complete.cases()`. Построчное удаление.
5. Простое восстановление данных с помощью пакета `Hmisc`.
6. Многомерное восстановление данных с помощью пакета `mice`. Моделирование набора массивов для пропущенных значений с помощью функции `mice()`. Применение необходимого статистического метода с помощью функции `with()` (например, линейная регрессия). Объединение результатов с помощью функции `pool()`.
7. Специализированные методы работы с пропущенными данными. Оценка методом максимального правдоподобия для многомерных нормальных данных с пропущенными значениями с помощью пакета `Mvnmle`.

Краткие методические указания

Проектное задание выполняется как во время практических занятий так и самостоятельно с использованием программ R-Studio и/или python. Преподаватель

излагает тему, приводит примеры и дает задание для самостоятельного выполнения. При необходимости, консультирует студентов.

Шкала оценки

№	Баллы	Описание
5	91-100	Студент выполнил проектное задание в полном соответствии с требованиями. Студент демонстрирует знания на высоком уровне. Демонстрирует способность создавать модели ИИ
4	76-90	Студент выполнил проектное задание не в полном соответствии с требованиями. Студент демонстрирует знания на среднем уровне. Демонстрирует способность использовать модели ИИ
3	61-75	Студент выполнил проектное задание частично. Студент демонстрирует знания на базовом уровне.
2	41-60	Студент демонстрирует знания на уровне ниже базового.
1	0-41	Студентом проявляется полное или практически полное отсутствие знаний и умений, но присутствовал на занятиях и пытался выполнять задание.

ФОС и ключи для ФОС
по дисциплине «Технологии сбора и предварительной обработки данных»

Критерий	Описание
Анализ формата данных в файле	Умение определить формат данных в файле и использовать стандартные средства для их загрузки
Анализ структуры данных в текстовом файле	Умение выявить структуру данных в произвольном текстовом файле и выбрать методы и средства для их загрузки
Умение программно загружать файлы из источников, расположенных в компьютерных сетях (локальных/глобальных)	Умение использовать вспомогательные программные средства для загрузки файлов из компьютерных сетей как локальных, так и из Интернет
Использование регулярных выражений для анализа и извлечения текстовых данных	Понимание принципов построения строковых шаблонов на языке регулярных выражений (RegExp). Умение применять программные инструментальные средства, реализующих функционал поиска, извлечения и замены строковых данных в тексте на основе регулярных выражений
Использование средств извлечения данных с веб-страниц	Умение использовать инструментарий для написания программ (на языке Python) для автоматизированного извлечения данных с веб-страниц. Умение использовать такие средства как BeautifulSoup и Selenium
Использование методов устранения пропусков и аномалий в данных	Умение обоснованно применять методы замены данных, основанные как на вероятностном подходе с использованием частот признаков, использование медианы, среднего или моды, так и с использованием более сложных методов, основанных на методе ближайших соседей и корреляции.

Шкала оценки для проектного задания

№	Баллы	Описание
5	91-100	Студент выполнил проектное задание в полном соответствии с требованиями. Студент демонстрирует знания на высоком уровне. Демонстрирует способность создавать модели ИИ
4	76-90	Студент выполнил проектное задание не в полном соответствии с требованиями. Студент демонстрирует знания на среднем уровне. Демонстрирует способность использовать модели ИИ
3	61-75	Студент выполнил проектное задание частично. Студент демонстрирует знания на базовом уровне.
2	41-61	Студент демонстрирует знания на уровне ниже базового.
1	0-40	Студентом проявляется полное или практически полное отсутствие знаний и умений, но присутствовал на занятиях и пытался выполнять

		задание.
--	--	----------