

МИНОБРНАУКИ РОССИИ



Федеральное государственное автономное образовательное учреждение

высшего образования

«Российский государственный гуманитарный университет»

(ФГАОУ ВО «РГГУ»)

ИНСТИТУТ ИНФОРМАЦИОННЫХ НАУК И ТЕХНОЛОГИЙ БЕЗОПАСНОСТИ

ФАКУЛЬТЕТ ИНФОРМАЦИОННЫХ СИСТЕМ И БЕЗОПАСНОСТИ

Кафедра информационных технологий и систем

Сбор и верификация данных

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

09.03.03 Прикладная информатика

Код и наименование направления подготовки/специальности

Прикладной искусственный интеллект

Наименование направленности (профиля)/ специализации

Уровень высшего образования: **бакалавр**

Форма обучения: **очная**

РПД адаптирована для лиц
с ограниченными возможностями
здоровья и инвалидов

Москва 2025

СБОР И ВЕРИФИКАЦИЯ ДАННЫХ

Рабочая программа дисциплины

Составитель:

к.т.н., доцент Д. Ю. Клехо

.....
.....

УТВЕРЖДЕНО

Протокол заседания кафедры ИТС

№6 от 12.12.2024

Оглавление

1. Пояснительная записка.....	4
1.1. Цель и задачи дисциплины.....	4
1.2. Формируемые компетенции, соотнесённые с планируемыми результатами обучения по дисциплине:.....	5
1.3. Место дисциплины в структуре образовательной программы.....	7
2. Структура дисциплины.....	8
3. Содержание дисциплины «Сбор и верификация данных».....	8
4. Образовательные технологии.....	8
5. Оценка планируемых результатов обучения.....	9
5.1 Система оценивания.....	9
5.2. Критерии выставления оценки по дисциплине.....	13
5.3. Оценочные средства для текущего контроля, промежуточной аттестации обучающихся.....	15
5.2. Описание контрольно-оценочных мероприятий промежуточного контроля по дисциплине.....	16
5.4 Содержание контрольно-оценочных мероприятий по воспитательной деятельности.....	16
6. Учебно-методическое и информационное обеспечение дисциплины.....	18
6.1 Список источников и литературы.....	18
6.2 Перечень ресурсов информационно-телекоммуникационной сети «Интернет».....	18
6.3 Профессиональные базы данных и информационно-справочные системы.....	19
7. Материально-техническое обеспечение дисциплины.....	19
8. Обеспечение образовательного процесса для лиц с ограниченными возможностями здоровья.....	19
9. Методические материалы.....	20
Планы практических занятий.....	20
Приложение 1.....	24
Аннотация рабочей программы дисциплины.....	24

1. Пояснительная записка

1.1. Цель и задачи дисциплины.

Дисциплина «Сбор и верификация данных» направлена на формирование у студентов компетенций в области сбора, подготовки и разметки данных для решения задач машинного обучения. От качества подобранных данных и их предварительной обработки зависит решение многих задач искусственного интеллекта. Студент в ходе обучения научится понимать какие процессы являются источниками данных, как данные создаются, как данные можно обрабатывать (алгоритмы и методы), как проводить предварительную подготовку данных, в каких форматах можно сохранять данные, как производить разметку и визуализацию данных для дальнейшего использования в задачах машинного обучения.

Задачи дисциплины:

1. Получение базовых знаний о процессе верификации и валидации данных, целей процесса верификации данных, месте верификации в жизненном цикле разработке программного обеспечения;
2. Приобретение теоретических знаний и практических навыков выбора и использования вычислительных систем для обработки информации на пользовательском уровне.
3. Ознакомление с основными моделями жизненного цикла программного обеспечения;
4. Освоение методов и средств разработки программного обеспечения;
5. Формирование практических навыков самостоятельного выявления, разработки, документирования, изменения и планирования требований с применением современных инструментальных средств.

1.2.Формируемые компетенции, соотнесённые с планируемыми результатами обучения по дисциплине:

Компетенции	Индикаторы достижения	Результаты обучения
2	3	4
ПК-1. Способен классифицировать и идентифицировать задачи искусственного интеллекта, выбирать адекватные методы и инструментальные средства решения задач искусственного интеллекта	ПК-1.1. Классифицирует и идентифицирует задачи систем искусственного интеллекта в зависимости от особенностей проблемной и предметной областей	<p><i>Знать:</i> основы математики, физики, вычислительной техники и программирования.</p> <p><i>Уметь:</i> решать стандартные профессиональные задачи с применением естественнонаучных и общетехнических знаний, методов математического анализа и моделирования.</p> <p><i>Владеть:</i> навыками в теоретическом и экспериментальном исследованиях объектов профессиональной деятельности.</p> <p><i>Знать:</i> стандартные методы решения задач теории</p>
	ПК-1.2. Выбирает методы и инструментальные средства искусственного интеллекта для решения задач в зависимости от особенностей проблемной и предметной областей	<p>вероятностей и математической статистики.</p> <p><i>Уметь:</i> решать стандартные профессиональные задачи с применением вероятностных методов, статистического анализа массовых явлений.</p> <p><i>Владеть:</i> навыками решения стандартных профессиональных задач с применением стандартных приемов прогноза событий и общепринятых таблиц классических стандартных распределений</p>

ПК-6. Способен осуществлять сбор и подготовку данных для систем искусственного интеллекта	ПК-6.1. Осуществляет поиск данных в открытых источниках, специализированных библиотеках и репозиториях	<p><i>Знать</i>: виды представления данных, методы поиска и парсинга данных, уровни представления данных (ODS, DDL, семантический слой, модель данных), основные инструменты, библиотеки и технологии Data Science</p> <p><i>Уметь</i>: отделять достоверные Источники данных от сомнительных, осуществлять критический отбор данных,</p>
	ПК-6.2. Выполняет подготовку и разметку структурированных и неструктурированных данных для машинного обучения	<p>проверять их на целостность непротиворечивость.</p> <p><i>Владеть</i>: инструментами и библиотеками Data Science для поиска данных в открытых источниках, специализированных библиотеках и репозиториях.</p> <p><i>Знать</i> основы экономической теории, системного анализа, математического моделирования.</p> <p><i>Уметь</i>: применять методы системного анализа, математического моделирования для анализа и разработки организационно-технических</p>
ПК-10: Способен решать задачи профессиональной деятельности в проектном формате для достижения заданной цели и создания уникального продукта, услуги или результата с заданным качеством в условиях ограниченности ресурсов, осознавая свою роль и ответственность в проекте	ПК-10.1 Решает задачи профессиональной деятельности в проектном формате для достижения заданной цели и создания уникального продукта, услуги или результата с заданным качеством	<p>и экономических процессов.</p> <p><i>Владеть</i>: навыками анализа и разработки организационно-технических и экономических процессов с применением методов системного анализа и математического моделирования.</p>
		<p><i>Знать</i> программно-аппаратные средства защиты информации в типовых операционных системах, системах управления базами данных, компьютерных сетях</p> <p><i>Уметь</i> конфигурировать программно-аппаратные средства защиты</p>

	ПК-10.2. Определяет возможные ограничения ресурсов в проектной деятельности	<p>информации в соответствии с заданными политиками безопасности</p> <p><i>Владеть</i> принципами формирования политики информационной безопасности объекта информатизации</p>
	ПК-10.3. Определяет свою роль и ответственность в команде при реализации проекта	<p><i>Знать</i> принципы формирования политики информационной безопасности в информационных системах</p> <p><i>Уметь</i> определять информационную инфраструктуру и информационные ресурсы организации, подлежащих защите;</p> <p><i>Владеть</i> навыками по разработке основных показателей технико-экономического обоснования соответствующих проектных решений</p> <p><i>Знать</i> основные этапы процесса проектирования и общие требования к содержанию проекта</p> <p><i>Уметь</i> анализировать показатели качества и критерии оценки систем и отдельных методов и средств защиты информации</p> <p><i>Владеть</i> принципами политики информационной безопасности объекта</p>

1.3. Место дисциплины в структуре образовательной программы

Дисциплина «Сбор и верификация данных» относится к обязательной части блока дисциплин учебного плана. Дисциплина «Сбор и верификация данных» входит в блок 1 «Дисциплины (модули)» образовательной программы 09.03.03 «Прикладная информатика». [2](#)

Дисциплина «Сбор и верификация данных» направлена на формирование у студентов компетенций в области сбора, подготовки и разметки данных для решения задач машинного обучения.

От качества подобранных данных и их предварительной обработки зависит решение многих задач искусственного интеллекта. Студент в ходе обучения

научится понимать какие процессы являются источниками данных, как данные создаются, как данные можно обрабатывать (алгоритмы и методы), как проводить предварительную подготовку данных, в каких форматах можно сохранять данные, как производить разметку и визуализацию данных для дальнейшего использования в задачах машинного обучения.

2. Структура дисциплины

Общая трудоёмкость дисциплины составляет 4 з.е., 144 академических часа.

Структура дисциплины для очной формы обучения

Объем дисциплины в форме контактной работы обучающихся с педагогическими работниками и (или) лицами, привлекаемыми к реализации образовательной программы на иных условиях, при проведении учебных занятий:

Семестр	Тип учебных занятий	Количество часов
5	Лекции	24
5	Практические занятия	32
Всего:		56

Объем дисциплины (модуля) в форме самостоятельной работы обучающихся составляет 70 академических часов.

3. Содержание дисциплины «Сбор и верификация данных».

Таблица 1.1

Код раздела, темы	Раздел, тема дисциплины	Содержание
1	Тема 1. Сбор и верификация данных	Ручной сбор данных. Парсинг данных. Создание данных. Математические модели, методы и алгоритмы сбора и верификации наукометрических данных.
2	Тема 2. Очистка данных	Поиск и устранение ошибок в данных. Заполнение недостающих данных. Качество данных. Форматы данных.
3	Тема 3. Маркировка данных	Разметка данных. Сценарии применения размеченных данных в задачах компьютерного зрения и обработки естественного языка.
4	Тема 4. Проверка и визуализация данных	Обработка данных. Закономерности и аномалии. Визуализация данных.

4. Образовательные технологии

Для проведения учебных занятий по дисциплине используются различные образовательные технологии. Для организации учебного процесса может быть использовано электронное обучение и (или) дистанционные образовательные технологии.

5. Оценка планируемых результатов обучения

5.1 Система оценивания

Форма контроля	Макс. количество баллов	
	За одну работу	Всего
Текущий контроль: - выполнение практических работ - контрольная работа в форме компьютерного тестирования	15 баллов 15 баллов	45 баллов 15 баллов
Промежуточная аттестация (Зачет)		40 баллов
Итого за семестр		100 баллов

Полученный совокупный результат конвертируется в традиционную шкалу оценок и в шкалу оценок Европейской системы переноса и накопления кредитов (European Credit Transfer System; далее – ECTS) в соответствии с таблицей:

100-балльная шкала	Традиционная шкала		Шкала ECTS
95 – 100	отлично	зачтено	A
83 – 94			B
68 – 82	хорошо		C
56 – 67	удовлетворительно		D
50 – 55			E
20 – 49	неудовлетворительно	не зачтено	FX
0 – 19			F

Баллы/ Шкала ECTS	Оценка по дисциплине	Критерии оценки результатов обучения по дисциплине
100-83/ A,B	отлично/ зачтено	<p>Выставляется обучающемуся, если он глубоко и прочно усвоил теоретический и практический материал, может продемонстрировать это на занятиях и в ходе промежуточной аттестации.</p> <p>Обучающийся исчерпывающе и логически стройно излагает учебный материал, умеет увязывать теорию с практикой, справляется с решением задач профессиональной направленности высокого уровня сложности, правильно обосновывает принятые решения.</p> <p>Свободно ориентируется в учебной и профессиональной литературе.</p> <p>Оценка по дисциплине выставляется обучающемуся с учётом результатов текущей и промежуточной аттестации.</p> <p>Компетенции, закреплённые за дисциплиной, сформированы на уровне – «высокий».</p>

Баллы/ Шкала ECTS	Оценка по дисциплине	Критерии оценки результатов обучения по дисциплине
82-68/ C	хорошо/ зачтено	<p>Выставляется обучающемуся, если он знает теоретический и практический материал, грамотно и по существу излагает его на занятиях и в ходе промежуточной аттестации, не допуская существенных неточностей.</p> <p>Обучающийся правильно применяет теоретические положения при решении практических задач профессиональной направленности разного уровня сложности, владеет необходимыми для этого навыками и приёмами.</p> <p>Достаточно хорошо ориентируется в учебной и профессиональной литературе.</p> <p>Оценка по дисциплине выставляется обучающемуся с учётом результатов текущей и промежуточной аттестации.</p> <p>Компетенции, закреплённые за дисциплиной, сформированы на уровне – «хороший».</p>
67-50/ D,E	удовлетвори- тельно/ зачтено	<p>Выставляется обучающемуся, если он знает на базовом уровне теоретический и практический материал, допускает отдельные ошибки при его изложении на занятиях и в ходе промежуточной аттестации.</p> <p>Обучающийся испытывает определённые затруднения в применении теоретических положений при решении практических задач профессиональной направленности стандартного уровня сложности, владеет необходимыми для этого базовыми навыками и приёмами.</p> <p>Демонстрирует достаточный уровень знания учебной литературы по дисциплине.</p> <p>Оценка по дисциплине выставляется обучающемуся с учётом результатов текущей и промежуточной аттестации.</p> <p>Компетенции, закреплённые за дисциплиной, сформированы на уровне – «достаточный».</p>
49-0/ F,FX	неудовлетворител- ьно/ не зачтено	<p>Выставляется обучающемуся, если он не знает на базовом уровне теоретический и практический материал, допускает грубые ошибки при его изложении на занятиях и в ходе промежуточной аттестации.</p> <p>Обучающийся испытывает серьёзные затруднения в применении теоретических положений при решении практических задач профессиональной направленности стандартного уровня сложности, не владеет необходимыми для этого навыками и приёмами.</p> <p>Демонстрирует фрагментарные знания учебной литературы по дисциплине.</p> <p>Оценка по дисциплине выставляется обучающемуся с учётом результатов текущей и промежуточной аттестации.</p>

Баллы/ Шкала ECTS	Оценка по дисциплине	Критерии оценки результатов обучения по дисциплине
		Компетенции на уровне «достаточный», закреплённые за дисциплиной, не сформированы.

Перечень ресурсов информационно-телекоммуникационной сети «Интернет».

Национальная электронная библиотека (НЭБ) www.rusneb.ru

ELibrary.ru Научная электронная библиотека www.elibrary.ru

Электронная библиотека Grebennikon.ru www.grebennikon.ru

Cambridge University Press

ProQuest Dissertation & Theses Global

SAGE Journals

Taylor and Francis

JSTOR

<http://znanium.com> – Электронно-библиотечная система.

<http://window.edu.ru> - Информационная система «Единое окно доступа к образовательным ресурсам».

Профессиональные базы данных и информационно-справочные системы

Доступ к профессиональным базам данных: <https://liber.rsuh.ru/ru/bases>

Информационные справочные системы:

Консультант Плюс

Гарант

Материально-техническое обеспечение дисциплины

Материально-техническое обеспечение дисциплины: Лаборатория аппаратных средств вычислительной техники – ауд. № 128:

№п/п	Оснащенность специальных помещений и помещений для самостоятельной работы	Перечень лицензионного программного обеспечения. Реквизиты подтверждающего документа		
		Наименование ПО	Лицензия/сертификат/заказ	Дата лицензии
1.	1 компьютер преподавателя, 10 компьютеров обучающихся, маркерная доска, проектор.	Windows 10 Microsoft office 2010 Pro Microsoft SQL Server 2008 Microsoft Visual Professional 2019 Mozilla Firefox 52.8.1 ESR CorelDrawCS6 Adobe CS4 Master Collection платформа ZOOM	68526624 49420326 46931055 63202190 свободный доступ 4097188 21375986	без даты 08.12.2011 20.05.2010 без даты свободный доступ 17.09.2012 13.01.2010

Обеспечение образовательного процесса для лиц с ограниченными возможностями здоровья

В ходе реализации дисциплины используются следующие дополнительные методы обучения, текущего контроля успеваемости и промежуточной аттестации обучающихся в зависимости от их индивидуальных особенностей:

- для слепых и слабовидящих: лекции оформляются в виде электронного документа, доступного с помощью компьютера со специализированным программным обеспечением; письменные задания выполняются на компьютере со специализированным программным обеспечением или могут быть заменены устным ответом; обеспечивается индивидуальное равномерное освещение не менее 300 люкс; для выполнения задания при необходимости предоставляется увеличивающее устройство; возможно также использование собственных увеличивающих устройств; письменные задания оформляются увеличенным шрифтом; экзамен и зачёт проводятся в устной форме или выполняются в письменной форме на компьютере.

- для глухих и слабослышащих: лекции оформляются в виде электронного документа, либо предоставляется звукоусиливающая аппаратура индивидуального пользования; письменные задания выполняются на компьютере в письменной форме; экзамен и зачёт проводятся в письменной форме на компьютере; возможно проведение в форме тестирования.

- для лиц с нарушениями опорно-двигательного аппарата: лекции оформляются в виде электронного документа, доступного с помощью компьютера со специализированным программным обеспечением; письменные задания выполняются на компьютере со специализированным программным обеспечением; экзамен и зачёт проводятся в устной форме или выполняются в письменной форме на компьютере.

При необходимости предусматривается увеличение времени для подготовки ответа.

Процедура проведения промежуточной аттестации для обучающихся устанавливается с учётом их индивидуальных психофизических особенностей. Промежуточная аттестация может проводиться в несколько этапов.

При проведении процедуры оценивания результатов обучения предусматривается использование технических средств, необходимых в связи с индивидуальными особенностями обучающихся. Эти средства могут быть предоставлены университетом, или могут использоваться собственные технические средства.

Проведение процедуры оценивания результатов обучения допускается с использованием дистанционных образовательных технологий.

Обеспечивается доступ к информационным и библиографическим ресурсам в сети Интернет для каждого обучающегося в формах, адаптированных к ограничениям их здоровья и восприятия информации:

- для слепых и слабовидящих: в печатной форме увеличенным шрифтом, в форме электронного документа, в форме аудиофайла.

- для глухих и слабослышащих: в печатной форме, в форме электронного документа.

- для обучающихся с нарушениями опорно-двигательного аппарата: в печатной форме, в форме электронного документа, в форме аудиофайла.

Учебные аудитории для всех видов контактной и самостоятельной работы, научная библиотека и иные помещения для обучения оснащены специальным оборудованием и учебными местами с техническими средствами обучения:

- для слепых и слабовидящих: устройством для сканирования и чтения с камерой SARA CE; дисплеем Брайля PAC Mate 20; принтером Брайля EmBraille ViewPlus;

- для глухих и слабослышащих: автоматизированным рабочим местом для людей с нарушением слуха и слабослышащих; акустический усилитель и колонки;

для обучающихся с нарушениями опорно-двигательного аппарата: передвижными, регулируемые эргономическими партами СИ-1; компьютерной техникой со специальным программным обеспечением.

Система оценивания (5 семестр)

Форма контроля	Макс. количество баллов	
	За одну	Всего

	работу	
Текущий контроль: - выполнение практических работ - контрольная работа в форме компьютерного тестирования	15 баллов 15 баллов	45 баллов 15 баллов
Промежуточная аттестация (Зачет)		40 баллов
Итого за семестр		100 баллов

Система оценивания (5 семестр).

Форма контроля	Макс. количество баллов	
	За одну работу	Всего
Текущий контроль: - выполнение практических работ - контрольная работа в форме компьютерного тестирования	15 баллов 15 баллов	45 баллов 15 баллов
Промежуточная аттестация (Экзамен)		40 баллов
Итого за семестр		100 баллов

Полученный совокупный результат конвертируется в традиционную шкалу оценок и в шкалу оценок Европейской системы переноса и накопления кредитов (European Credit Transfer System; далее – ECTS) в соответствии с таблицей:

100-балльная шкала	Традиционная шкала		Шкала ECTS
95 – 100	отлично	зачтено	A
83 – 94			B
68 – 82	хорошо		C
56 – 67	удовлетворительно		D
50 – 55		E	
20 – 49	неудовлетворительно	не зачтено	FX
0 – 19			F

5.2. Критерии выставления оценки по дисциплине

Баллы/ Шкала ECTS	Оценка по дисциплине	Критерии оценки результатов обучения по дисциплине
100-83/ A,B	отлично/ зачтено	Выставляется обучающемуся, если он глубоко и прочно усвоил теоретический и практический материал, может продемонстрировать это на занятиях и в ходе промежуточной аттестации. Обучающийся исчерпывающе и логически стройно излагает учебный материал, умеет увязывать теорию с практикой, справляется с решением задач профессиональной направленности высокого уровня сложности, правильно обосновывает принятые решения. Свободно ориентируется в учебной и профессиональной литературе.

Баллы/ Шкала ECTS	Оценка по дисциплине	Критерии оценки результатов обучения по дисциплине
		<p>Оценка по дисциплине выставляются обучающемуся с учётом результатов текущей и промежуточной аттестации.</p> <p>Компетенции, закреплённые за дисциплиной, сформированы на уровне – «высокий».</p>
82-68/ С	хорошо/ зачтено	<p>Выставляется обучающемуся, если он знает теоретический и практический материал, грамотно и по существу излагает его на занятиях и в ходе промежуточной аттестации, не допуская существенных неточностей.</p> <p>Обучающийся правильно применяет теоретические положения при решении практических задач профессиональной направленности разного уровня сложности, владеет необходимыми для этого навыками и приёмами.</p> <p>Достаточно хорошо ориентируется в учебной и профессиональной литературе.</p> <p>Оценка по дисциплине выставляются обучающемуся с учётом результатов текущей и промежуточной аттестации.</p> <p>Компетенции, закреплённые за дисциплиной, сформированы на уровне – «хороший».</p>
67-50/ D,E	удовлетвори- тельно/ зачтено	<p>Выставляется обучающемуся, если он знает на базовом уровне теоретический и практический материал, допускает отдельные ошибки при его изложении на занятиях и в ходе промежуточной аттестации.</p> <p>Обучающийся испытывает определённые затруднения в применении теоретических положений при решении практических задач профессиональной направленности стандартного уровня сложности, владеет необходимыми для этого базовыми навыками и приёмами.</p> <p>Демонстрирует достаточный уровень знания учебной литературы по дисциплине.</p> <p>Оценка по дисциплине выставляются обучающемуся с учётом результатов текущей и промежуточной аттестации.</p> <p>Компетенции, закреплённые за дисциплиной, сформированы на уровне – «достаточный».</p>
49-0/ F,FX	неудовлетворител- ьно/ не зачтено	<p>Выставляется обучающемуся, если он не знает на базовом уровне теоретический и практический материал, допускает грубые ошибки при его изложении на занятиях и в ходе промежуточной аттестации.</p> <p>Обучающийся испытывает серьёзные затруднения в применении теоретических положений при решении практических задач профессиональной направленности стандартного уровня сложности, не владеет</p>

Баллы/ Шкала ECTS	Оценка по дисциплине	Критерии оценки результатов обучения по дисциплине
		необходимыми для этого навыками и приёмами. Демонстрирует фрагментарные знания учебной литературы по дисциплине. Оценка по дисциплине выставляются обучающемуся с учётом результатов текущей и промежуточной аттестации. Компетенции на уровне «достаточный», закреплённые за дисциплиной, не сформированы.

5.3. Оценочные средства для текущего контроля, промежуточной аттестации обучающихся.

Примерные вопросы на компьютерное тестирование:

ОПК – 2: ОПК-2.1, ОПК-2.2, ОПК-2.3

5.1 Описание аудиторных контрольно-оценочных мероприятий по дисциплине модуля

5.1.1. Лекции

Самостоятельное изучение теоретического материала по темам/разделам лекций в соответствии с содержанием дисциплины (п. 1.2. РПД)

5.1.2. Лабораторные занятия

Примерный перечень тем

1. Аннотирование изображений ограничивающими прямоугольниками
2. Аннотирование изображений многоугольниками
3. Семантическая сегментация
4. 3D-кубоиды
5. Аннотирование ключевыми точками
6. Аннотирование видео
7. Категоризация текста
8. Аннотирование аудио
9. Генерация

синтетических данных

LMS-платформа – не предусмотрена

5.2 Описание внеаудиторных контрольно-оценочных мероприятий и средств текущего контроля по дисциплине модуля

Разноуровневое (дифференцированное) обучение.

Базовый

Домашняя работа

Примерный перечень тем

1. Методы обработки и сегментации изображений

Примерные задания:

1. Ознакомьтесь с содержанием предложенных изображений

2. Разработайте программу, осуществляющую:
 - 2.1 Ввод изображений
 - 2.2 Сегментацию изображений на основе точечных и пространственных преобразований
 - 2.3 Генерацию признаков описаний объектов
 - 2.4 Классификацию объектов
 3. Подготовьте отчет о проделанной работе.
- LMS-платформа – не предусмотрена

5.2. Описание контрольно-оценочных мероприятий промежуточного контроля по дисциплине

Зачет

Список примерных вопросов

1. Ручной и автоматизированный сбор данных.
2. Парсинг данных. Создание синтетических данных.
3. Очистка данных. Поиск и устранение ошибок в данных.
4. Качество данных. Форматы данных.
5. Маркировка данных.
6. Разметка изображений.
7. Семантическая сегментация.
8. Категоризация текста.
9. Разметка текста.
10. Метрики для измерения сходств данных.
11. Биометрические данные и их классификация.
12. Анализ данных о формах объектов на изображениях.

LMS-платформа – не предусмотрена

5.4 Содержание контрольно-оценочных мероприятий по воспитательной деятельности

Направлени е воспитательн ой деятельности	Вид воспитательно й деятельности	Технология воспитательной деятельности	Компетен ция	Результат ы обучения	Контрольно- оценочные мероприят ия
Профессиона льное воспитание	профориента ционная деятельность	Технология самостоятельн ой работы Технология анализа образовательн ых задач	ПК-6	ПК-6.1. У-2	Домашняя работа Зачет Лабораторн ые занятия Лекции

5.4.1. Лабораторные занятия

Примерный перечень тем

1. Аннотирование изображений ограничивающими прямоугольниками
 2. Аннотирование изображений многоугольниками
 3. Семантическая сегментация
 4. 3D-кубоиды
 5. Аннотирование ключевыми точками
 6. Аннотирование видео
 7. Категоризация текста
 8. Аннотирование аудио
 9. Генерация синтетических данных LMS-платформа – не предусмотрена
- 5.4.2. Описание внеаудиторных контрольно-оценочных мероприятий и средств текущего контроля по дисциплине модуля**

**Разноуровневое (дифференцированное) обучение.
Базовый**

5.4.3. Домашняя работа

Примерный перечень тем

Методы обработки и сегментации изображений.

Примерные задания:

1. Ознакомьтесь с содержанием предложенных изображений
2. Разработайте программу, осуществляющую:
 - 2.1 Ввод изображений
 - 2.2 Сегментацию изображений на основе точечных и пространственных преобразований
 - 2.3 Генерацию признаков описаний объектов
 - 2.4 Классификацию объектов
3. Подготовьте отчет о проделанной работе.
4. LMS-платформа – не предусмотрена

5.5. Описание контрольно-оценочных мероприятий промежуточного контроля по дисциплине модуля

5.5.1. Зачет

Список примерных вопросов

1. Ручной и автоматизированный сбор данных.
2. Парсинг данных. Создание синтетических данных.
3. Очистка данных. Поиск и устранение ошибок в данных.
4. Качество данных. Форматы данных.
5. Маркировка данных.
6. Разметка изображений.
7. Семантическая сегментация.
8. Категоризация текста.
9. Разметка текста.
10. Метрики для измерения сходств данных.
11. Биометрические данные и их классификация.
12. Анализ данных о формах объектов на изображениях.
13. LMS-платформа – не предусмотрена

6. Учебно-методическое и информационное обеспечение дисциплины

6.1 Список источников и литературы

Основная

1. Лаврищева, Е. М. Программная инженерия и технологии программирования сложных систем : учебник для вузов / Е. М. Лаврищева. — 2-е изд., испр. и доп. — Москва : Издательство Юрайт, 2021 — 432 с. — (Высшее образование). — ISBN 978-5-534-07604-2. — Текст : электронный // ЭБС Юрайт [сайт]. — URL: <https://urait.ru/bcode/470923>
2. Галушка В.В. Формирование обучающей выборки при использовании искусственных нейронных сетей в задачах поиска ошибок баз данных [Текст] / В.В. Галушка, В.А. Фатхи // Инженерный Вестник Допа [Электронный ресурс]: электрон. науч.-инновац. журн. — 2013. — №2. — Режим доступа: <http://ivdon.ru/mag3zine/archive/n2y2013/1597>. — Загл. с экрана. — № гос. регистрации 0421100096.
3. Adrion, W., Branstad, M. and Cherniavski, “Validation, verification and testing of computer software”, – Adrion, 1982
4. Гаврилова Т.А., Хорошевский В.Ф. Базы знаний интеллектуальных систем. Учебник / Гаврилова Т.А. [и др.]; – СПб. : Изд-во «Питер», 2001
5. Сеницын, С. В. Верификация программного обеспечения : краткий учебный курс / С. В. Сеницын, Н. Ю. Налютин. - Москва : ИНТУИТ, 2016. - 320 с. - ISBN 978-5-94774-825-3. - Текст : электронный. - URL: <https://znanium.ru/catalog/product/2138810> (дата обращения: 30.01.2025). – Режим доступа: по подписке.

Дополнительная литература

1. Сысоева, Л. А. Управление проектами информационных систем : учебное пособие / Л.А. Сысоева, А.Е. Сатунина. — Москва : ИНФРА-М, 2021 — 345 с.— DOI 10.12737/textbook_5cc01bbf923e13.56817630. - ISBN 978-5-16-013775-9. - Текст : электронный. - URL: <https://znanium.com/catalog/product/1167942>
2. Рыбина Г.В. Автоматизированное построение баз знаний для интегрированных экспертных систем // Изв. РАН. Теория и системы управления. № 5,1998, С.152-166.
3. Галушка В.В. Исследование моделей дефектов искусственных нейронных сетей [Текст] / В.В. Галушка, Д.В. Маршаков, В.А. Фатхи // Вестник Дон. гос. техн. ун-та. — 2012. — №3 (64). — стр. 65-71.
4. Аверина, Т. А. Численные методы. Верификация алгоритмов решения систем со случайной структурой : учебное пособие для вузов / Т. А. Аверина. — Москва : Издательство Юрайт, 2020 — 179 с. — (Высшее образование). — ISBN 978-5-534-07205-1. — Текст : электронный // ЭБС Юрайт [сайт]. — URL: <https://urait.ru/bcode/455466>

6.2 Перечень ресурсов информационно-телекоммуникационной сети

«Интернет».

Национальная электронная библиотека (НЭБ) www.rusneb.ru

ELibrary.ru Научная электронная библиотека www.elibrary.ru

Электронная библиотека Grebennikon.ru www.grebennikon.ru

Cambridge University Press

ProQuest Dissertation & Theses Global

SAGE Journals

Taylor and Francis

JSTOR

<http://znanium.com> – Электронно-библиотечная система.

<http://window.edu.ru> - Информационная система «Единое окно доступа к образовательным ресурсам».

6.3 Профессиональные базы данных и информационно-справочные системы

Доступ к профессиональным базам данных: <https://liber.rsuh.ru/ru/bases>

Информационные справочные системы:

1. Консультант Плюс
2. Гарант

7. Материально-техническое обеспечение дисциплины

Материально-техническое обеспечение дисциплины: Лаборатория аппаратных средств вычислительной техники – ауд. № 128:

№п/п	Оснащенность специальных помещений и помещений для самостоятельной работы	Перечень лицензионного программного обеспечения. Реквизиты подтверждающего документа		
		Наименование ПО	Лицензия/сертификат/заказ	Дата лицензии
2.	1 компьютер преподавателя, 10 компьютеров обучающихся, маркерная доска, проектор.	Windows 10 Microsoft office 2010 Pro Microsoft SQL Server 2008 Microsoft Visual Professional 2019 Mozilla Firefox 52.8.1 ESR CorelDrawCS6 Adobe CS4 Master Collection платформа ZOOM	68526624 49420326 46931055 63202190 свободный доступ 4097188 21375986	без даты 08.12.2011 20.05.2010 без даты свободный доступ 17.09.2012 13.01.2010

8. Обеспечение образовательного процесса для лиц с ограниченными возможностями здоровья.

В ходе реализации дисциплины используются следующие дополнительные методы обучения, текущего контроля успеваемости и промежуточной аттестации обучающихся в зависимости от их индивидуальных особенностей: для слепых и слабовидящих: лекции оформляются в виде электронного документа, доступного с помощью компьютера со специализированным программным обеспечением; письменные задания выполняются на компьютере со специализированным программным обеспечением или могут быть заменены устным ответом; обеспечивается индивидуальное равномерное освещение не менее 300 люкс; для выполнения задания при необходимости предоставляется увеличивающее устройство; возможно также использование собственных увеличивающих устройств; письменные задания оформляются увеличенным шрифтом; экзамен и зачёт проводятся в устной форме или выполняются в письменной форме на компьютере.

- для глухих и слабослышащих: лекции оформляются в виде электронного документа, либо предоставляется звукоусиливающая аппаратура индивидуального пользования; письменные задания выполняются на компьютере в письменной форме; экзамен и зачёт проводятся в письменной форме на компьютере; возможно проведение в форме тестирования.

- для лиц с нарушениями опорно-двигательного аппарата: лекции оформляются в виде электронного документа, доступного с помощью компьютера со специализированным программным обеспечением; письменные задания выполняются на компьютере со специализированным программным обеспечением; экзамен и зачёт проводятся в устной форме или выполняются в письменной форме на компьютере.

При необходимости предусматривается увеличение времени для подготовки ответа.

Процедура проведения промежуточной аттестации для обучающихся устанавливается с учётом их индивидуальных психофизических особенностей. Промежуточная аттестация может проводиться в несколько этапов.

При проведении процедуры оценивания результатов обучения предусматривается использование технических средств, необходимых в связи с индивидуальными особенностями обучающихся. Эти средства могут быть предоставлены университетом, или могут использоваться собственные технические средства.

Проведение процедуры оценивания результатов обучения допускается с использованием дистанционных образовательных технологий.

Обеспечивается доступ к информационным и библиографическим ресурсам в сети Интернет для каждого обучающегося в формах, адаптированных к ограничениям их здоровья и восприятия информации:

- для слепых и слабовидящих: в печатной форме увеличенным шрифтом, в форме электронного документа, в форме аудиофайла.
- для глухих и слабослышащих: в печатной форме, в форме электронного документа.
- для обучающихся с нарушениями опорно-двигательного аппарата: в печатной форме, в форме электронного документа, в форме аудиофайла.

Учебные аудитории для всех видов контактной и самостоятельной работы, научная библиотека и иные помещения для обучения оснащены специальным оборудованием и учебными местами с техническими средствами обучения:

- для слепых и слабовидящих: устройством для сканирования и чтения с камерой SARA CE; дисплеем Брайля PAC Mate 20; принтером Брайля EmBraille ViewPlus;
- для глухих и слабослышащих: автоматизированным рабочим местом для людей с нарушением слуха и слабослышащих; акустический усилитель и колонки;
- для обучающихся с нарушениями опорно-двигательного аппарата: передвижными, регулируемые эргономическими партами СИ-1; компьютерной техникой со специальным программным обеспечением.

9. Методические материалы

Планы практических занятий.

Работа с пропущенными данными.

- 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 максимального правдоподобия для многомерных нормальных данных с пропущенными значениями с помощью пакета **Mvnmle**.

Тестовые задания:

@Верификация это...

- + процесс проверки соответствия поведения системы требованиям
- процесс устранения ошибок в программном обеспечении
- процесс взаимодействия с пользователем, направленный на улучшение его понимания принципов работы программной системы
- процесс уточнения требований по результатам обсуждения с пользователем

@Дайте определение тестирования, как вида деятельности...

- это процесс доказательства того, что программная система соответствует ожиданиям пользователя или заказчика
- это процесс доказательства того, что программная реализация системы и требования на систему соответствуют друг другу и проектным стандартам
- + это процесс поиска и документирования дефектов программной реализации разрабатываемой системы
- это процесс поиска и исправления ошибок в проектной документации и программной реализации системы

@Дайте определение верификации, как вида деятельности...

- это процесс поиска и документирования дефектов программной @реализации разрабатываемой системы
- + это процесс доказательства того, что программная реализация системы и требования на систему соответствуют друг другу и проектным стандартам
- это процесс доказательства того, что программная система соответствует ожиданиям пользователя или заказчика
- это процесс поиска и исправления ошибок в проектной документации и программной реализации системы

@Укажите основные отличия тест-планов от тест-требований

- тест-планы служат для создания тестовых сценариев
- + тест-планы описывают конкретные способы тестирования системы
- тест-планы описывают общие подходы к тестированию
- тест-планы пишутся на основе функциональных требований

@Какую цель имеет процедура идентификации объектов конфигурации?

- + присвоение каждому объекту конфигурации уникального имени, позволяющего отличать его от других
- определение процедуры идентификации пользователей для доступа к объектам

- определение схемы аудита идентифицированных объектов конфигурации
- предотвращение неверных правил доступа к объектам конфигурации

@Какую цель имеет процедура идентификации объектов конфигурации?

- доступность исходного текста в качестве справочного материала для тестировщика
- полное сокрытие исходного текста и определение входных и выходных значений для тестов не только из требований
- + полное сокрытие исходного текста от тестировщика и определение входных и выходных значений для тестов только из требований
- полное отсутствие возможности проверить наличие недеklarированного поведения системы

@Модульное тестирование предназначено для...

- функционирования одного замкнутого участка программного модуля
- проверки функционирования каждого независимого программного модуля
- + для тестирования модуля в условиях отсутствия воздействия со стороны пользователя
- для максимальной изоляции побочного функционирования модуля со стороны остальных частей системы.

@Укажите основные достоинства восходящего интеграционного тестирования

- усложняется локализация дефектов в модулях
- упрощается процедура разработки тестового окружения
- уменьшается количество работ по сборке модулей
- + упрощается локализация дефектов в модулях и дефекты выявляются на ранних стадиях тестирования

@Какова цель процедуры управления изменениями?

- обеспечение возможности хранить изменения независимо от данных проекта
- + обеспечения контроля управляемости изменений, вносимых в систему и целостности системы после внесения изменений
- создание отчетов для аудита процесса управления конфигурациями
- поддержка процесса архивации

@Каким свойством не обладает требование "Проверить, что система работает правильно"?

- полноты
- + тестируемости
- непротиворечивости
- замкнутости

@Самый большой этап в жизненном цикле программы:

- изучение предметной области;

- программирование;
- + корректировка ошибок;
- эксплуатация;

@Защитное программирование это:

- встраивание в программу отладочных средств;
- создание задач, защищенных от копирования;
- разделение доступа в программе;
- + использование паролей;

@Отладка – это:

- правило вызова процедур (функций);
- процедура поиска ошибок, когда известно, что ошибка есть;
- составление блок-схемы алгоритма.
- + определение списка параметров;

@Когда программист может проследить последовательность выполнения команд программы:

- при трассировке;
- + при тестировании;
- при компиляции;
- при выполнении программы.

@Что такое оптимизация программ:

- улучшение работы существующей программы;
- разработка модульной конструкции программы;
- + создание удобного интерфейса пользователя;
- применение методов объектно-ориентированного программирования.

Аннотация рабочей программы дисциплины

Дисциплина «Сбор и верификация данных» направлена на формирование у студентов компетенций в области сбора, подготовки и разметки данных для решения задач машинного обучения. От качества подобранных данных и их предварительной обработки зависит решение многих задач искусственного интеллекта. Студент в ходе обучения научится понимать какие процессы являются источниками данных, как данные создаются, как данные можно обрабатывать (алгоритмы и методы), как проводить предварительную подготовку данных, в каких форматах можно сохранять данные, как производить разметку и визуализацию данных для дальнейшего использования в задачах машинного обучения.

Задачи дисциплины:

1. Получение базовых знаний о процессе верификации и валидации данных, целей процесса верификации данных, месте верификации в жизненном цикле разработке программного обеспечения;
2. Приобретение теоретических знаний и практических навыков выбора и использования вычислительных систем для обработки информации на пользовательском уровне.
3. Ознакомление с основными моделями жизненного цикла программного обеспечения;
4. Освоение методов и средств разработки программного обеспечения;
5. Формирование практических навыков самостоятельного выявления, разработки, документирования, изменения и планирования требований с применением современных инструментальных средств.

Знать: методы сбора и обобщения информации о проблемной области путем опроса экспертов, сходных данных о функционировании проблемной и предметной областей, документированных источников знаний, а также формирования требований к системе искусственного интеллекта виды представления данных, методы поиска и парсинга данных, уровни представления данных(ODS, DDL, семантический слой, модель данных), основные инструменты, библиотеки и технологии Data Science.

Уметь: отделять достоверные источники данных от сомнительных, осуществлять критический отбор данных, проверять их на целостность непротиворечивость, осуществлять сбор и обобщение информации о проблемной области путем опроса экспертов, исходных данных о функционировании проблемной области, документированных источников знаний, а также формировать требования к системе искусственного интеллекта

Владеть: навыками в теоретическом и экспериментальном исследованиях объектов профессиональной деятельности. инструментами и библиотеками Data Science для поиска данных в открытых источниках, специализированных библиотеках и репозиториях.