Документ подписан простой электронной подписью
Информация о владельце:
ФИО: Максимов Алексей Борисович
Должность: директор департамента по образовательной политике
Дата подписания: 04.10.2023 16:37:51
Уникальный программный ключ:
8db180d1a3f02ac9e60521a5672742735c18b1d6

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ

«МОСКОВСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»

(МОСКОВСКИЙ ПОЛИТЕХ)

Факультет машиностроения

УТВЕРЖДАЮ

Декан

_____ /Е.В. Сафонов/

«27» апреля 2023 г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

Интеллектуальный анализ данных

Направление подготовки
11.03.01 Радиотехника

Профиль
Интеллектуальная радиоэлектроника и промышленный интернет вещей

Квалификация
Бакалавр

Формы обучения
очная

Москва, 2023 г.

Разработчик(и):

Старший преподаватель кафедры АиУ  /Т.А. Лисовская/

Согласовано:
Заведующий кафедрой «Автоматика и управление»,
д.т.н., профессор  /А.А. Радионов/

Руководитель образовательной программы
д.т.н., профессор  /А.А. Радионов/

Содержание

1 Цели, задачи и планируемые результаты обучения по дисциплине 4
2 Место дисциплины в структуре образовательной программы 4
3 Структура и содержание дисциплины 5
3.1 Виды учебной работы и трудоемкость 5
3.2 Тематический план изучения дисциплины 5
3.3 Содержание дисциплины 6
3.4 Тематика семинарских/практических и лабораторных занятий 6
3.5 Тематика курсовых проектов (курсовых работ) 7
4 Учебно-методическое и информационное обеспечение 7
4.1 Нормативные документы и ГОСТы 7
4.2 Основная литература 7
4.3 Дополнительная литература 7
4.4 Электронные образовательные ресурсы 8
4.5 Лицензионное и свободно распространяемое программное обеспечение 8
4.6 Современные профессиональные базы данных и информационные справочные системы 8
5 Материально-техническое обеспечение 8
6 Методические рекомендации 8
6.1 Методические рекомендации для преподавателя по организации обучения 8
6.2 Методические указания для обучающихся по освоению дисциплины 9
7 Фонд оценочных средств 10
7.1 Методы контроля и оценивания результатов обучения 10
7.2 Шкала и критерии оценивания результатов обучения 12
7.3 Оценочные средства 15

1 Цели, задачи и планируемые результаты обучения по дисциплине

Целью изучения дисциплины "Интеллектуальный анализ данных" состоит в развитии у студентов представления о методах обработки, хранения и представления данных, об их структуризации, выявлении тенденций. Изучаются алгоритмы математической статистики и классического машинного обучения и их применение в интеллектуальных системах в промышленности.

Задачи изучения дисциплины:

- ознакомление студентов с задачами, принципами, методами и подходами обработки данных;
- приобретение теоретических знаний, практических умений и навыков в области исследования задач анализа данных и их решения методами машинного обучения.

Обучение по дисциплине «Интеллектуальный анализ данных» направлено на формирование у обучающихся следующих компетенций:

Код и наименование компетенций	Индикаторы достижения компетенции	Наименование показателя оценивания
ПК-7. Способен разрабатывать и применять методы искусственного интеллекта и машинного обучения для решения профессиональных задач	ИПК-7.1 Выбирает методы и инструментальные средства искусственного интеллекта и машинного обучения для решения задач в зависимости от особенностей проблемной и предметной областей ИПК-7.2 Определяет метрики оценки результатов моделирования и критерии качества построенных моделей ИПК-7.3 Принимает участие в оценке, выборе и при необходимости разработке методов искусственного интеллекта и машинного обучения	**Знать:** Методы математической статистики и анализа данных для создания моделей систем искусственного интеллекта для решения профессиональных задач **Уметь:** Выбирать и применять алгоритмы анализа данных и оценивать показатели качества построенных моделей **Владеть:** Навыками написания программного кода основных алгоритмов математической статистики, анализа и визуализации данных, а также машинного обучения

2 Место дисциплины в структуре образовательной программы

Дисциплина относится к части, формируемой участниками образовательных отношений блока Б1 «Дисциплины (модули)».

Дисциплина непосредственно связана со следующими дисциплинами и практиками ООП:

Информационные технологии;

Компьютерное зрение;

Машинное обучение;
Объектно-ориентированное программирование;
Основы алгоритмизации и программирования.

3 Структура и содержание дисциплины

Общая трудоемкость дисциплины составляет 3 зачетные единицы (108 часов).

3.1 Виды учебной работы и трудоемкость

№ п/п	Вид учебной работы	Количество часов	Семестр
			6
1	**Аудиторные занятия**	**54**	**54**
	В том числе:		
1.1	Лекции	18	18
1.2	Семинарские/практические занятия	36	36
2	**Самостоятельная работа**	**54**	**54**
	В том числе:		
2.1	Подготовка отчётов по практическим работам	24	24
2.2	Работа с конспектом лекций	16	16
2.3	Подготовка к диф.зачёту	14	14
3	**Промежуточная аттестация**		
	Зачет/диф.зачет/экзамен		диф.зачет
	Итого	**108**	**108**

3.2 Тематический план изучения дисциплины

№ п/п	Разделы/темы дисциплины	Трудоемкость, час					
		Всего	Аудиторная работа				Самостоятельная работа
			Лекции	Семинарские/ практические занятия	Лабораторные занятия	Практическая подготовка	
1	**Раздел 1.**	**24**	**6**	**12**	**0**	**0**	**6**
1.1	Тема 1. Введение в интеллектуальный анализ данных: основные понятия и методология		2	4			2
1.2	Тема 2. Основы математической статистики для анализа данных		2	4			2
1.3	Тема 3. Основы машинного обучения: типы задач и основные подходы.		2	4			2
2	**Раздел 2.**	**56**	**8**	**16**	**0**	**0**	**32**
2.1	Тема 1. Обучение с учителем: задача классификации и методы.		2	4			8
2.2	Тема 2. Обучение с учителем: задача регрессии и методы.		2	4			8

2.3	Тема 3. Обучение без учителя: кластеризация и методы.		2	4			8
2.4	Тема 4. Обучение без учителя: снижение размерности и анализ данных.		2	4			8
3	**Раздел 3**	**28**	**4**	**8**	**0**	**0**	**16**
3.1	Тема 1. Оценка моделей и переобучение: методы кросс-валидации и регуляризации.		2	4			8
3.2	Тема 2. Применение интеллектуального анализа данных в реальных проектах.		2	4			8
	Итого	**108**	**18**	**36**	**0**	**0**	**54**

3.3 Содержание дисциплины

Раздел 1. Введение. Математическая статистика для обработки информации

Призван ознакомить с основами интеллектуального анализа данных. В данном разделе рассматриваются ключевые понятия и методология, связанные с этой областью. Основное внимание уделяется типам данных, методам их получения, хранения, сортировки и т.п. Раздел также включает в себя введение в интеллектуальный анализ данных, обсуждение основ классической математической статистики, необходимых для анализа данных, и ознакомление с основными типами задач и подходами в машинном обучении.

Раздел 2. Обучение с учителем и без учителя.

В разделе приводятся подробности обучения моделей, на основе форматированных данных. Здесь подробно исследуются методы обучения с учителем, такие как задачи классификации и регрессии. Также освещается обучение без учителя, включая методы кластеризации для выявления паттернов и структур в данных. Методы рассматриваются как с классической точки зрения математической статистики, так и с точки зрения современных основ машинного обучения.

Раздел 3. Интеллектуальный анализ данных в реальных задачах.

Углубляется в практическое применение анализа данных. Здесь обсуждаются методы оценки моделей и преодоление проблемы переобучения с помощью техник кросс-валидации и регуляризации. Также акцент делается на том, как интеллектуальный анализ данных находит свое применение в реальных проектах, включая задачи снижения размерности, анализа данных и извлечения практических выводов из больших объемов информации.

3.4 Тематика семинарских/практических и лабораторных занятий

3.4.1 Семинарские/практические занятия

Семинар 1. Введение в интеллектуальный анализ данных: основные концепции и цели анализа данных; роль интеллектуального анализа данных в современном мире;

Семинар 2. Основы математической статистики для анализа данных

Семинар 3. Знакомство с средой и инструментами для анализа данных.

Семинар 4. Основы работы с данными: загрузка, предобработка, визуализация.

Семинар 5. Применение статистических методов к анализу данных: распределения, корреляции.

Семинар 6-7. Классификация данных: логистическая регрессия и метод опорных векторов.

Семинар 8. Классификация данных: методы деревьев решений.

Семинар 9-10. Регрессионный анализ: линейная и полиномиальная регрессия.

Семинар 11. Кластерный анализ: методы и интерпретация результатов.

Семинар 12-13. Методы кластеризации: k-средних, иерархическая кластеризация и другие

Семинар 14-15. Снижение размерности данных: метод главных компонент и t-SNE.

Семинар 16. Методы оценки моделей: кросс-валидация, метрики качества.

Семинар 17-18. Практическое применение интеллектуального анализа данных

3.4.2 Лабораторные занятия

Не предусмотрены

3.5 Тематика курсовых проектов (курсовых работ)

Не предусмотрены

4 Учебно-методическое и информационное обеспечение

4.1 Нормативные документы и ГОСТы

Не предусмотрены

4.2 Основная литература

1. Коэльо, Л. П. Построение систем машинного обучения на языке Python / Л. П. Коэльо, В. Ричарт ; перевод с английского А. А. Слинкин. — 2-е изд. — Москва : ДМК Пресс, 2016. — 302 с. — ISBN 978-5-97060-330-7. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/82818

2. Миркин, Б. Г. Введение в анализ данных: учебник и практикум / Б. Г. Миркин. — Москва : Издательство Юрайт, 2023. — 174 с. — (Высшее образование). — ISBN 978-5-9916-5009-0. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: https://urait.ru/bcode/511121

3. Мхитарян, В.С. Анализ данных: учебник для вузов / В. С. Мхитарян [и др.]; под редакцией В. С. Мхитаряна. — Москва: Издательство Юрайт, 2023. — 490 с. — (Высшее образование). — ISBN 978-5-534-00616-2. — Текст: электронный // Образовательная платформа Юрайт [сайт]. — URL: https://urait.ru/bcode/511020

4.3 Дополнительная литература

1. Рашка, С. Python и машинное обучение: крайне необходимое пособие по новейшей предсказательной аналитике, обязательное для более глубокого понимания методологии машинного обучения : руководство / С. Рашка ; перевод с английского А. В. Логунова. — Москва : ДМК Пресс, 2017. — 418 с. — ISBN 978-5-97060-409-0. — Текст : электронный // Лань : электронно-библиотечная система. — URL: https://e.lanbook.com/book/100905

2. Анализ научно-технических данных и результатов исследований: учебник для вузов / А. Н. Асаул, Е. И. Рыбнов, Г. Ф. Щербина, М. А. Асаул. — Москва: Издательство Юрайт, 2023. — 240 с. — (Высшее образование). — ISBN 978-5-534-15448-1. — Текст: электронный // Образовательная платформа Юрайт [сайт]. — URL: https://urait.ru/bcode/520339

3. Мойзес, Б.Б. Статистические методы контроля качества и обработка экспериментальных данных: учебное пособие для вузов / Б. Б. Мойзес, И. В. Плотникова, Л. А. Редько. — 2-е изд. — Москва: Издательство Юрайт, 2022; Томск: Томский политехнический университет. — 118 с. — (Высшее образование). — ISBN 978-5-534-11906-0 (Издательство Юрайт). — ISBN 978-5-4387-0700-4 (Томский политехнический университет). — Текст: электронный // Образовательная платформа Юрайт [сайт]. — URL: https://urait.ru/bcode/495895

4.4 Электронные образовательные ресурсы

Не предусмотрены

4.5 Лицензионное и свободно распространяемое программное обеспечение

1. Anaconda
2. Python

4.6 Современные профессиональные базы данных и информационные справочные системы

1. Научная электронная библиотека http://www.elibrary.ru
2. Российская государственная библиотека http://www.rsl.ru
3. Единое окно доступа к образовательным ресурсам Федеральный портал http://window.edu.ru
4. ЭБС «Университетская библиотека онлайн» https://biblioclub.ru/index.php

5 Материально-техническое обеспечение

1. Компьютерный класс с предустановленным программным обеспечением, указанным в п. 4.5, мультимедийное оборудование (проектор, персональный компьютер преподавателя).
2. Аудитория для лекционных, практических занятий. Оборудование и аппаратура: аудиторная доска, возможность использования мультимедийного комплекса.

6 Методические рекомендации

6.1 Методические рекомендации для преподавателя по организации обучения

На первом занятии по дисциплине необходимо ознакомить студентов с порядком ее изучения (темами курса, формами занятий, текущего и промежуточного контроля), раскрыть место и роль дисциплины в системе наук, ее практическое значение, довести до студентов требования к форме отчетности и применения видов контроля.

При подготовке к практическим работам по перечню объявленных тем преподавателю необходимо уточнить план их проведения, продумать формулировки и содержание учебных вопросов, выносимых на обсуждение.

В ходе работы во вступительном слове раскрыть практическую значимость темы работы, определить порядок ее проведения, время на обсуждение каждого учебного вопроса.

Применяя фронтальный опрос дать возможность выступить всем студентам, присутствующим на занятии.

В заключительной части работы следует подвести ее итоги: дать оценку выступлений каждого студента и учебной группы в целом. Раскрыть положительные стороны и недостатки проведенной работы. Ответить на вопросы студентов. Выдать задания для самостоятельной работы по подготовке к следующему занятию.

Методика преподавания дисциплины «Интеллектуальный анализ данных» и реализация компетентностного подхода в изложении и восприятии материала предусматривает использование следующих активных и интерактивных форм проведения групповых, индивидуальных, аудиторных занятий в сочетании с внеаудиторной работой с целью формирования и развития профессиональных навыков обучающихся:

– подготовка к выполнению и защита практических работ с помощью специализированного программного обеспечения;

– технологии анализа ситуаций для активного обучения, которые позволяют студентам соединить теорию и практику, представить примеры принимаемых решений и их последствий, демонстрировать различные позиции, формировать навыки оценки альтернативных вариантов в вероятностных условиях.

Обучение по дисциплине ведется с применением традиционных потоково-групповых информационно-телекоммуникационных технологий. При осуществлении образовательного процесса по дисциплине используются следующие информационно-телекоммуникационные технологии: презентации с применением проектора и программы PowerPoint.

6.2 Методические указания для обучающихся по освоению дисциплины

Самостоятельная работа является одним из видов учебных занятий. Цель самостоятельной работы – практическое самостоятельное получение студентами навыков работы в программных продуктах, рассматриваемых в процессе изучения дисциплины.

Аудиторная самостоятельная работа по дисциплине выполняется на учебных занятиях под непосредственным руководством преподавателя и по его заданию.

Внеаудиторная самостоятельная работа выполняется студентом по заданию преподавателя, но без его непосредственного участия.

Задачи самостоятельной работы студента:

- развитие навыков самостоятельной учебной работы;
- освоение содержания дисциплины;
- углубление содержания и осознание основных понятий дисциплины;
- использование материала, собранного и полученного в ходе самостоятельных занятий для эффективной подготовки к зачёту.

Виды внеаудиторной самостоятельной работы:

- самостоятельное изучение отдельных тем дисциплины;
- подготовка к практическим занятиям;
- оформление отчетов по выполненным практическим работам;

Для выполнения любого вида самостоятельной работы необходимо пройти следующие этапы:

- определение цели самостоятельной работы;
- конкретизация познавательной задачи;
- самооценка готовности к самостоятельной работе;
- выбор адекватного способа действия, ведущего к решению задачи;
- планирование работы (самостоятельной или с помощью преподавателя) над заданием;
- осуществление в процессе выполнения самостоятельной работы самоконтроля (промежуточного и конечного) результатов работы, и корректировка выполнения работы;
- рефлексия;

- презентация работы.

7 Фонд оценочных средств

В процессе обучения используются следующие оценочные формы самостоятельной работы студентов, оценочные средства текущего контроля успеваемости и промежуточных аттестаций

- отчёты по практическим работам;
- контрольные работы;
- дифференцированный зачёт.

Оценочные средства текущего контроля успеваемости включают контрольные задания по практическим работам индивидуально для каждого обучающегося.

В результате освоения дисциплины (модуля) формируются следующие компетенции:

Код компетенции	**Наименование компетенции выпускника**
ПК-7	Способен разрабатывать и применять методы искусственного интеллекта и машинного обучения для решения профессиональных задач

7.1 Методы контроля и оценивания результатов обучения

Перечень оценочных средств по дисциплине «Интеллектуальный анализ данных»

№ п/п	Вид контроля результатов обучения	Наименование контроля результатов обучения	Краткая характеристика контроля результатов обучения
	Текущий	Практическая работа	Практическая работа выполняется индивидуально каждым студентом. По результатам выполнения работы студент оформляет отчёт, содержащий подробное описание проделанной работы с наглядным представлением результатов. Оформленный отчет студент сдает преподавателю на проверку в заранее установленный срок. При проверке преподаватель оценивает качество оформления, корректность проделанных шагов, результатов работы и выводов. Часть работ подразумевает устную защиту в формате доклада/презентации.
	Текущий	Контрольная работа	Контрольная работа выполняется индивидуально каждым студентом. Задание подразумевает выполнения ряда шагов с использованием

			программного обеспечения, изучаемого в соответствующей теме. Результатом выполнения контрольной работы является электронный документ с кратким описанием выполненных шагов и полученных результатов. При проверке преподаватель оценивает как результат работы, так и пусть достижения результата.
	Промежуточный	Дифференцированный зачёт	Промежуточная аттестация обучающихся в форме дифференцированного зачёта проводится по результатам выполнения всех видов учебной работы, предусмотренных учебным планом по данной дисциплине (модулю), при этом учитываются результаты текущего контроля успеваемости в течение семестра. Оценка степени достижения обучающимися планируемых результатов обучения по дисциплине (модулю) проводится преподавателем, ведущим занятия по дисциплине (модулю) методом экспертной оценки. По итогам промежуточной аттестации по дисциплине (модулю) выставляется оценка «отлично», «хорошо», «удовлетворительно» или «неудовлетворительно». Во время проведения дифференцированного зачёта студенту предлагается выполнить 3 практических задания. По результату выполнения студенту задаются вопросы на понимание выполненных задач. Количество дополнительных вопросов зависит от полноты ответа студента. К промежуточной аттестации допускаются только студенты, выполнившие все виды учебной работы, предусмотренные рабочей программой по дисциплине «Интеллектуальный анализ данных»

7.2 Шкала и критерии оценивания результатов обучения

Показателем оценивания компетенций на различных этапах их формирования является достижение обучающимися планируемых результатов обучения по дисциплине (модулю).

Показатель	Критерии оценивания			
	2	**3**	**4**	**5**
знать: Методы математической статистики и анализа данных для создания моделей систем искусственного интеллекта для решения профессиональных задач	Обучающийся демонстрирует полное отсутствие или недостаточное соответствие знаний: методы математической статистики и анализа данных для создания моделей систем искусственного интеллекта для решения профессиональных задач	Обучающийся демонстрирует неполное соответствие знаний: методы математической статистики и анализа данных для создания моделей систем искусственного интеллекта для решения профессиональных задач. Допускаются значительные ошибки, проявляется недостаточность знаний, по ряду показателей, обучающийся испытывает значительные затруднения при оперировании знаниями при их переносе на новые ситуации.	Обучающийся демонстрирует частичное соответствие знаний: методы математической статистики и анализа данных для создания моделей систем искусственного интеллекта для решения профессиональных задач. Допускаются незначительные ошибки, неточности, затруднения при аналитических операциях.	Обучающийся демонстрирует полное соответствие знаний: методы математической статистики и анализа данных для создания моделей систем искусственного интеллекта для решения профессиональных задач. Свободно оперирует приобретенными знаниями.
уметь: Выбирать и применять алгоритмы анализа данных и оценивать показатели качества построенных моделей	Обучающийся не умеет или в недостаточной степени умеет выбирать выбирать и применять алгоритмы анализа данных и оценивать показатели качества построенных моделей	Обучающийся демонстрирует неполное соответствие умений выбирать и применять алгоритмы анализа данных и оценивать показатели качества построенных моделей. Допускаются значительные ошибки, проявляется недостаточность умений, по ряду показателей, обучающийся испытывает значительные затруднения при оперировании умениями при их	Обучающийся демонстрирует частичное соответствие умений выбирать и применять алгоритмы анализа данных и оценивать показатели качества построенных моделей. Умения освоены, но допускаются незначительные ошибки, неточности, затруднения при аналитических операциях, переносе умений на новые, нестандартные ситуации.	Обучающийся демонстрирует полное соответствие умений выбирать и применять алгоритмы анализа данных и оценивать показатели качества построенных моделей. Свободно оперирует приобретенными умениями, применяет их в ситуациях повышенной сложности.

		переносе на новые ситуации.		
владеть: Навыками написания программного кода основных алгоритмов математической статистики, анализа и визуализации данных, а также машинного обучения.	Обучающийся не владеет или в недостаточной степени владеет навыками написания программного кода основных алгоритмов математической статистики, анализа и визуализации данных, а также машинного обучения.	Обучающийся в недостаточной степени владеет: навыками написания программного кода основных алгоритмов математической статистики, анализа и визуализации данных, а также машинного обучения. Обучающийся испытывает значительные затруднения при применении навыков в новых ситуациях.	Обучающийся частично владеет: навыками написания программного кода основных алгоритмов математической статистики, анализа и визуализации данных, а также машинного обучения. Навыки освоены, но допускаются незначительные ошибки, неточности, затруднения при аналитических операциях, переносе умений на новые, нестандартные ситуации.	Обучающийся в полном объеме владеет: навыками написания программного кода основных алгоритмов математической статистики, анализа и визуализации данных, а также машинного обучения. Свободно применяет полученные навыки в ситуациях повышенной сложности.

Шкала оценивания промежуточной аттестации: дифференцированный зачёт

Шкала оценивания	Описание
Отлично	Выполнены все виды учебной работы, предусмотренные учебным планом. Студент демонстрирует соответствие знаний, умений, навыков приведенным в таблицах показателям, оперирует приобретенными знаниями, умениями, навыками, применяет их в ситуациях повышенной сложности, не испытывает затруднений при аналитических операциях, переносе знаний и умений на новые, нестандартные ситуации.
Хорошо	Выполнены все виды учебной работы, предусмотренные учебным планом. Студент демонстрирует частичное соответствие знаний, умений, навыков приведенным в таблицах показателям, оперирует приобретенными знаниями, умениями, навыками, применяет их в ситуациях повышенной сложности. При этом могут быть допущены незначительные ошибки, неточности, затруднения при аналитических операциях, переносе знаний и умений на новые, нестандартные ситуации.
Удовлетворительно	Выполнены все виды учебной работы, предусмотренные учебным планом. Студент демонстрирует неполное соответствие знаний, умений, навыков приведенным в таблицах показателям, допускаются значительные ошибки, проявляется отсутствие знаний, умений, навыков по ряду показателей, студент испытывает значительные затруднения при оперировании знаниями и умениями при их переносе на новые ситуации.
Неудовлетворительно	Не выполнен один или более видов учебной работы, предусмотренных учебным планом. Студент демонстрирует полное отсутствие или недостаточное соответствие знаний, умений, навыков приведенным в таблицах показателям,

	допускаются значительные ошибки, проявляется отсутствие знаний, умений, навыков по ряду показателей, студент не может оперировать знаниями и умениями при их переносе на новые ситуации.

Шкала оценивания текущего контроля

Наименование контроля результатов обучения	Шкала оценивания	Описание
Практическая работа	Зачтено: набрано 2 и более баллов. Не зачтено: набрано 1 и менее баллов. Критерии оценивания: Общий балл при оценке складывается из следующих показателей: - оформление работы соответствует требованиям – 1 балл; - все пункты задания выполнены в полном объёме – 2 балла; - изложение, описание и выводы по работе грамотны и полно описывают содержание практической работы – 2 балла.	В качестве форм текущего контроля знаний студентов используются отчеты по практической работе Отчёт по практической работе содержит описания ряда шагов по выполнению практической работы согласно заданию с подробным описанием проделанных действий и полученными результатами. Защита практических работ (если требуется) осуществляется индивидуально. Студентом предоставляется оформленный отчёт. Студенты, не выполнившие практическую работу, к защите не допускаются
Контрольная работа	Отлично - Работа высокого качества, уровень выполнения отвечает всем требованиям, теоретическое содержание курса освоено полностью, без пробелов, необходимые практические навыки работы с освоенным материалом сформированы, все предусмотренные программой обучения учебные задания выполнены, либо некоторые из выполненных заданий содержат незначительные ошибки Хорошо - Уровень выполнения работы отвечает большинству основных требований, теоретическое содержание курса освоено частично, но пробелы не носят существенного	Защита темы включает выполнения практического задания по изученному материалу в аудитории в течении одной пары. Билеты состоят из задач, позволяющих оценить сформированные компетенции. На решение отводится 1.5 часа.

	характера, необходимые практические навыки работы с освоенным материалом в основном сформированы, некоторые виды заданий выполнены с ошибками. Удовлетворительно - Теоретическое содержание курса освоено частично, необходимые практические навыки работы не сформированы, большинство предусмотренных программой заданий не выполнено; при дополнительной самостоятельной работе над материалом курса возможно повышение качества выполнения учебных заданий. Неудовлетворительно - Теоретическое содержание курса не освоено, необходимые практические навыки работы не сформированы, предусмотренные программой задания не выполнены	

7.3 Оценочные средства

7.3.1 Текущий контроль

Типовое задание «Практическая работа 1»

Задание 1: Напишите Python-скрипт для загрузки данных из CSV-файла, выполните базовую статистическую предобработку (средние, медианы, стандартные отклонения) и выведите основные характеристики данных.

Задание 2: На основе случайной выборки данных, рассчитайте выборочное среднее и дисперсию. Затем создайте график гистограммы распределения данных.

Типовое задание «Практическая работа 2»

Задание 1. Используя библиотеку Pandas, загрузите данные о продажах из Excel-файла. Выполните фильтрацию, сортировку и группировку данных, а затем выведите сводную статистику.

Задание 2. Загрузите набор данных о пассажирах Титаника. Проведите предварительную очистку (удаление пропущенных значений) и визуализацию (гистограммы возрастов, столбчатая диаграмма классов).

Типовое задание «Практическая работа 3»

Задание 1. Рассчитайте корреляцию между двумя переменными из набора данных и визуализируйте их взаимосвязь с помощью диаграммы рассеяния.

Задание 2. Напишите Python-скрипты для реализации логистической регрессии и метода опорных векторов. Используйте библиотеки NumPy и Scikit-learn. Протестируйте модели на наборе данных и выведите метрики классификации (точность, полнота, F1-мера).

Типовое задание «Практическая работа 4»

Задание 1. Примените метод k-средних для кластеризации данных. Оцените число оптимальных кластеров с помощью "метода локтя" и визуализируйте полученные кластеры.

Задание 2. Сравните результаты кластеризации, полученные методами k-средних и иерархической кластеризации. Визуализируйте дендрограмму для иерархической кластеризации.

Типовое задание «Практическая работа 5»

Задание 1. Примените метод главных компонент и t-SNE для снижения размерности данных. Визуализируйте данные в новом пространстве и сравните результаты методов.

Типовое задание «Практическая работа 6»

Задание. Примените кросс-валидацию для оценки качества модели. Рассчитайте различные метрики классификации/регрессии для разных моделей и сравните их результаты.

Типовое задание к контрольной работе 1.

Темы: Введение в интеллектуальный анализ данных, Основы математической статистики для анализа данных

1. Опишите роль интеллектуального анализа данных в современном мире и приведите примеры его практического применения.

2. Рассчитайте среднее, медиану и стандартное отклонение для данного набора данных: [12, 15, 18, 20, 25, 30]. Визуализируйте распределение данных с помощью гистограммы.

Типовое задание к контрольной работе 2.

Темы: Знакомство с средой и инструментами для анализа данных, Основы работы с данными: загрузка, предобработка, визуализация

1. Загрузите данные о погоде из CSV-файла в Pandas DataFrame. Очистите данные от дубликатов и пропущенных значений.

2. Визуализируйте среднюю температуру по месяцам в виде линейного графика. Добавьте подписи **осей и легенду.**

Типовое задание к контрольной работе 3.

Темы: Применение статистических методов к анализу данных, Классификация данных: логистическая регрессия и метод опорных векторов

1. Рассчитайте коэффициент корреляции между двумя переменными в наборе данных. Оцените его силу и направление связи.

2. Используя библиотеку Scikit-learn, обучите модели логистической регрессии и метода опорных векторов на наборе данных. Оцените точность классификации для каждой модели.

Типовое задание к контрольной работе 4.

Темы: Классификация данных: методы деревьев решений, Регрессионный анализ: линейная и полиномиальная регрессия

1. Постройте решающее дерево для набора данных о предметах и их характеристиках. Визуализируйте полученное дерево и проанализируйте полученные разбиения.

2. С использованием данных о продажах автомобилей, постройте линейную и полиномиальную регрессии для прогнозирования стоимости автомобиля. Оцените точность моделей и выберите наилучшую.

7.3.2 Промежуточная аттестация

Какие основные задачи решает интеллектуальный анализ данных?	ПК-7
Что такое предобработка данных и какие этапы включает?	ПК-7
Какие основные методы визуализации данных вы знаете?	ПК-7
Чем отличается средняя арифметическая от медианы?	ПК-7
Что такое дисперсия и зачем она используется?	ПК-7
Какие математические статистики помогают описать данные?	ПК-7
Как можно загрузить данные в Python с использованием библиотеки Pandas?	ПК-7
Что такое корреляция и каковы её типы?	ПК-7
Для чего используется метод k-средних в кластерном анализе?	ПК-7
Какой метод используется для снижения размерности данных?	ПК-7
Какова цель кросс-валидации при оценке моделей?	ПК-7

Какие метрики используются для оценки качества моделей классификации?	ПК-7
Что такое решающее дерево и как оно строится?	ПК-7
Какова основная идея логистической регрессии?	ПК-7
Какие этапы включает процесс регрессионного анализа?	ПК-7
Что такое переобучение и как его избежать?	ПК-7
Какие шаги необходимы для кластеризации данных методом иерархической кластеризации?	ПК-7
Что такое метод опорных векторов и в чём заключается его принцип?	ПК-7
Каким образом осуществляется предварительная обработка данных перед классификацией?	ПК-7
Какие методы применяются для оценки важности признаков в моделях?	ПК-7
Что такое метрика F1-мера и для чего она используется?	ПК-7
Какие основные шаги включает процесс анализа данных?	ПК-7
Какие типы графиков подходят для визуализации категориальных данных?	ПК-7
Какие типы распределений вы знаете и как они выглядят?	ПК-7
Какие методы снижения размерности данных помогают сохранить важные характеристики?	ПК-7
Что такое p-значение в контексте статистического анализа?	ПК-7
Какие этапы включает процесс построения решающего дерева?	ПК-7
Какие параметры влияют на обобщающую способность модели?	ПК-7
Какой алгоритм используется для определения оптимального числа кластеров в методе k-средних?	ПК-7
Какие основные этапы включает процесс классификации данных?	ПК-7
Как можно оценить качество кластеризации без заранее известных меток?	ПК-7
Что такое коэффициент детерминации R^2 в регрессионном анализе?	ПК-7
Какие метрики качества используются для оценки результатов кластеризации?	ПК-7
Как выбрать наилучшую модель классификации с помощью кросс-валидации?	ПК-7
Что такое гиперпараметры модели и как они влияют на её обучение?	ПК-7
Как можно обработать выбросы в данных перед анализом?	ПК-7
Какие методы можно использовать для борьбы с несбалансированными классами в классификации?	ПК-7
Что такое мультиколлинеарность в регрессионном анализе?	ПК-7
Какие проблемы могут возникнуть при переобучении модели?	ПК-7
Что такое PCA и как он используется для снижения размерности данных?	ПК-7
Какие шаги включает процесс применения алгоритма кластеризации?	ПК-7
Что такое confusion matrix и какие метрики могут быть вычислены на её основе?	ПК-7
Каким образом можно настроить параметры решающего дерева, чтобы избежать переобучения?	ПК-7
Какие методы используются для заполнения пропущенных значений в данных?	ПК-7
Что такое ROC-кривая и какие метрики можно получить на её основе?	ПК-7
Какие типы задач классификации вы знаете?	ПК-7
Что такое высокая дисперсия модели и какие могут быть причины её возникновения?	ПК-7
Какие методы визуализации могут помочь понять структуру кластеров?	ПК-7
Какой алгоритм используется для определения оптимального числа кластеров в иерархической кластеризации?	ПК-7
Какие методы могут использоваться для создания ансамблей моделей классификации?	ПК-7