

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ
ФЕДЕРАЦИИ
федеральное государственное автономное образовательное учреждение высшего
образования
"САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
АЭРОКОСМИЧЕСКОГО ПРИБОРОСТРОЕНИЯ"

Кафедра № 32

УТВЕРЖДАЮ

Руководитель направления

доц., к.т.н., доц.

(должность, уч. степень, звание)

С.В. Солёный

(инициалы, фамилия)



(подпись)

«22» июня 2023 г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

«Машинное обучение и анализ данных»
(Наименование дисциплины)

Код направления подготовки	13.04.02
Наименование направления подготовки	Электроэнергетика и электротехника
Наименование направленности	Цифровая энергетика
Форма обучения	очная

Лист согласования рабочей программы дисциплины

Программу составил (а)

проф., д.т.н., доц.
(должность, уч. степень, звание)


(подпись, дата)

С.А. Сериков
(инициалы, фамилия)

Программа одобрена на заседании кафедры № 32
«24» апреля 2023 г., протокол № 6

Заведующий кафедрой № 32

доц., к.т.н., доц.
(уч. степень, звание)


(подпись, дата)

С.В. Солёный
(инициалы, фамилия)

Ответственный за ОП ВО 13.04.02(03)

доц., к.т.н., доц.
(должность, уч. степень, звание)


(подпись, дата)

О.Я. Солёная
(инициалы, фамилия)

Заместитель директора института №3 по методической работе

старший преподаватель
(должность, уч. степень, звание)


(подпись, дата)

Н.В. Решетникова
(инициалы, фамилия)

Аннотация

Дисциплина «Машинное обучение и анализ данных» входит в образовательную программу высшего образования – программу магистратуры по направлению подготовки 13.04.02 «Электроэнергетика и электротехника» направленности «Цифровая энергетика». Дисциплина реализуется кафедрой «№32».

Дисциплина нацелена на формирование у выпускника следующих компетенций:

УК-1 «Способен осуществлять критический анализ проблемных ситуаций на основе системного подхода, выработать стратегию действий»

ОПК-2 «Способен применять современные методы исследования, оценивать и представлять результаты выполненной работы»

Содержание дисциплины охватывает круг вопросов, связанных с применением методов машинного обучения и интеллектуального анализа данных при решении прикладных задач в области электроэнергетики и электротехники.

Преподавание дисциплины предусматривает следующие формы организации учебного процесса: лекции, практические занятия, самостоятельная работа обучающегося.

Программой дисциплины предусмотрены следующие виды контроля: текущий контроль успеваемости, промежуточная аттестация в форме дифференцированного зачета.

Общая трудоемкость освоения дисциплины составляет 3 зачетных единицы, 108 часов.

Язык обучения по дисциплине «русский».

1. Перечень планируемых результатов обучения по дисциплине

1.1. Цель преподавания дисциплины: формирование у студентов комплекса профессиональных знаний и умений в области интеллектуального анализа данных и машинного обучения.

1.2. Дисциплина входит в состав обязательной части образовательной программы высшего образования (далее – ОП ВО).

1.3. Перечень планируемых результатов обучения по дисциплине, соотнесенных с планируемыми результатами освоения ОП ВО.

В результате изучения дисциплины обучающийся должен обладать следующими компетенциями или их частями. Компетенции и индикаторы их достижения приведены в таблице 1.

Таблица 1 – Перечень компетенций и индикаторов их достижения

Категория (группа) компетенции	Код и наименование компетенции	Код и наименование индикатора достижения компетенции
Универсальные компетенции	УК-1 Способен осуществлять критический анализ проблемных ситуаций на основе системного подхода, выработать стратегию действий	УК-1.Д.1 выявляет причинно-следственные связи и анализирует объект как систему УК-1.Д.2 осуществляет поиск вариантов решения поставленной проблемной ситуации на основе доступных источников информации и выработывает стратегию действий для решения проблемной ситуации УК-1.Д.3 умеет анализировать, сохранять и передавать информацию с использованием цифровых средств
Общепрофессиональные компетенции	ОПК-2 Способен применять современные методы исследования, оценивать и представлять результаты выполненной работы	ОПК-2.Д.2 выбирает и применяет современные средства представления и обработки данных, использует алгоритмы машинного обучения

2. Место дисциплины в структуре ОП

Дисциплина может базироваться на знаниях, ранее приобретенных обучающимися при изучении следующих дисциплин:

- «Математические методы и модели в научных исследованиях»,
- «Электрические системы и сети».

Знания, полученные при изучении материала данной дисциплины, имеют как самостоятельное значение, так и могут использоваться при изучении других дисциплин:

- «Интеллектуальные системы энергосбережения»,
- «Цифровые двойники в электроэнергетике»,
- «Электрические системы и сети».

3. Объем и трудоемкость дисциплины

Данные об общем объеме дисциплины, трудоемкости отдельных видов учебной работы по дисциплине (и распределение этой трудоемкости по семестрам) представлены в таблице 2.

Таблица 2 – Объем и трудоемкость дисциплины

Вид учебной работы	Всего	Трудоемкость по семестрам
		№1
1	2	3
Общая трудоемкость дисциплины, ЗЕ/ (час)	3/ 108	3/ 108
Из них часов практической подготовки		
Аудиторные занятия, всего час.	34	34
в том числе:		
лекции (Л), (час)	17	17
практические/семинарские занятия (ПЗ), (час)	17	17
лабораторные работы (ЛР), (час)		
курсовой проект (работа) (КП, КР), (час)		
экзамен, (час)		
Самостоятельная работа, всего (час)	74	74
Вид промежуточной аттестации: зачет, дифф. зачет, экзамен (Зачет, Дифф. зач, Экз.**)	Дифф. Зач.	Дифф. Зач.

Примечание: ** кандидатский экзамен

4. Содержание дисциплины

4.1. Распределение трудоемкости дисциплины по разделам и видам занятий. Разделы, темы дисциплины и их трудоемкость приведены в таблице 3.

Таблица 3 – Разделы, темы дисциплины, их трудоемкость

Разделы, темы дисциплины	Лекции (час)	ПЗ (СЗ) (час)	ЛР (час)	КП (час)	СРС (час)
Семестр 1					
Введение	2	0			5
Раздел 1. Основные понятия машинного обучения	2	4			15
Раздел 2. Линейные методы классификации	4	4			10
Раздел 3. Метрики качества классификации	3	3			10
Раздел 4. Решающие деревья	2	2			14
Раздел 5. Композиция алгоритмов	2	2			15
Заключение	2	2			5
Итого в семестре:	17	17			74
Итого	17	17	0	0	74

Практическая подготовка заключается в непосредственном выполнении обучающимися определенных трудовых функций, связанных с будущей профессиональной деятельностью.

4.2. Содержание разделов и тем лекционных занятий.

Содержание разделов и тем лекционных занятий приведено в таблице 4.

Таблица 4 – Содержание разделов и тем лекционного цикла

Номер раздела	Название и содержание разделов и тем лекционных занятий
1	Введение Предмет дисциплины и ее задачи. Структура и содержание дисциплины. Взаимосвязь изучаемой дисциплины с другими дисциплинами учебного плана. Характеристика основной и дополнительной литературы
2	Основные понятия машинного обучения Общая характеристика методов и разделов, входящих в предметную область - машинное обучение. Область применения и задачи, решаемые технологиями машинного обучения, основные понятия и обозначения. Основные библиотеки языка Python для работы с данными
3	Линейные методы классификации Метод стохастического градиента для настройки линейных классификаторов, регуляризация, метод опорных векторов, логистическая регрессия, линейная регрессия и общие работы с линейными методами
4	Метрики качества классификации Разновидности метрик качества бинарной и многоклассовой классификации, способы сведения многоклассовых задач к двухклассовым, прикладная интерпретация различных метрик.
5	Решающие деревья Подходы к отбору признаков, применение метода главных компонент, для понижения размерности.
6	Композиция алгоритмов Композиции деревьев и случайные леса, основные понятия и постановки задач, связанные с композициями, градиентный бустинг.
7	Заключение Основные тенденции развития технологий машинного обучения

4.3. Практические (семинарские) занятия

Темы практических занятий и их трудоемкость приведены в таблице 5.

Таблица 5 – Практические занятия и их трудоемкость

№ п/п	Темы практических занятий	Формы практических занятий	Трудоемкость, (час)	Из них практической подготовки, (час)	№ раздела дисциплины
Семестр 1					
1	Базовые фреймворки языка Python: знакомство с основами языка программирования Python при работе в среде Spider, Jupyter, numpy, matplotlib, pandas, seaborn.	групповая	4		2, 3
2	Решение задачи регрессии:	групповая	4		4

	построение модели линейной регрессии с использованием библиотеки scikit-learn, метрики				
3	Решение задачи классификации: недообучение, переобучение, логистическая регрессия, метрики	групповая	3		5
4	Алгоритмы машинного обучения: К- ближайших соседей, деревья решений	групповая	4		6
5	Алгоритмы машинного обучения: бутстрап, случайный лес, бустинг, стэкинг. Дополнительные методы обработки данных: работа с текстом, байесовский классификатор, предобработка данных.	групповая	2		7
Всего			17		

4.4. Лабораторные занятия

Темы лабораторных занятий и их трудоемкость приведены в таблице 6.

Таблица 6 – Лабораторные занятия и их трудоемкость

№ п/п	Наименование лабораторных работ	Трудоемкость, (час)	Из них практической подготовки, (час)	№ раздела дисциплины
Учебным планом не предусмотрено				
Всего				

4.5. Курсовое проектирование/ выполнение курсовой работы

Учебным планом не предусмотрено

4.6. Самостоятельная работа обучающихся

Виды самостоятельной работы и ее трудоемкость приведены в таблице 7.

Таблица 7 – Виды самостоятельной работы и ее трудоемкость

Вид самостоятельной работы	Всего, час	Семестр 1, час
1	2	3
Изучение теоретического материала дисциплины (ТО)	50	50
Курсовое проектирование (КП, КР)		
Расчетно-графические задания (РГЗ)		

Выполнение реферата (Р)		
Подготовка к текущему контролю успеваемости (ТКУ)	10	10
Домашнее задание (ДЗ)		
Контрольные работы заочников (КРЗ)		
Подготовка к промежуточной аттестации (ПА)	14	14
Всего:	74	74

5. Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине (модулю)
Учебно-методические материалы для самостоятельной работы обучающихся указаны в п.п. 7-11.

6. Перечень печатных и электронных учебных изданий
Перечень печатных и электронных учебных изданий приведен в таблице 8.
Таблица 8– Перечень печатных и электронных учебных изданий

Шифр/ URL адрес	Библиографическая ссылка	Количество экземпляров в библиотеке (кроме электронных экземпляров)
УДК 681.3.07	Рашка Себастьян, Мирджалили Вахид. Python и машинное обучение: машинное и глубокое обучение с использованием Python, scikit-learn и TensorFlow 2, 3-е изд. Пер. с англ. – СПб.: «Диалектика», 2020. – 848 с.	
УДК: 681.3.07	Жерон Орельен. Прикладное машинное обучение с помощью Scikit-Learn, Keras и TensorFlow: концепции, инструменты и техники для создания интеллектуальных систем, 2-е изд.: Пер. с англ.– СПб.: ООО «Диалектика», 2020.–1040 с.	
УДК 004.43 ПЗ7	Плас Дж. Вандер. Python для сложных задач: наука о данных и машинное обучение. – СПб.: Питер, 2018. – 576 с.	
УДК 681.3.07	Джоши Прадик. Искусственный интеллект с примерами на Python.: Пер. с англ. - СПб.: ООО "Диалектика", 2019. – 448 с.	
ISBN 978-5- 9908910-8-1	Андреас Мюллер, Сара Гвидо. Введение в машинное обучение с помощью python. Руководство для специалистов по работе с данными. – М.: ИЦ «Гевисста», 2017. – 393 с.	
УДК 0004.438	Луис Педро Коэльо, Вилли Ричарт. Построение систем машинного обучения на языке Python. 2-е издание/ пер. с англ. Слинкин А.А. – М.: ДМК Пресс, 2016. – 302 с.	

УДК 681.3.07	Элбон Крис. Машинное обучение с использованием Python. Сборник рецептов: Пер. с англ. — СПб.: БХВ-Петербург, 2019. — 384 с.: ил.	
УДК 004.62	Силен Дэви, Мейсман Арно, Али Мохамед. Основы Data Science и Big Data. Python и науки о данных. – СПб.: Питер, 2017. – 336 с.	

7. Перечень электронных образовательных ресурсов информационно-телекоммуникационной сети «Интернет»

Перечень электронных образовательных ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины приведен в таблице 9.

Таблица 9 – Перечень электронных образовательных ресурсов информационно-телекоммуникационной сети «Интернет»

URL адрес	Наименование
http://www.machinelearning.ru/wiki/index.php?title=Заглавная_страница	Профессиональный информационно-аналитический ресурс, посвященный машинному обучению, распознаванию образов и интеллектуальному анализу данных
http://python.org	Официальный сайт языка Python
https://www.anaconda.com/download/	Anaconda - Дистрибутив Python,

8. Перечень информационных технологий

8.1. Перечень программного обеспечения, используемого при осуществлении образовательного процесса по дисциплине.

Перечень используемого программного обеспечения представлен в таблице 10.

Таблица 10– Перечень программного обеспечения

№ п/п	Наименование
1	Anaconda 2020.11 For Windows Installer
2	Python 3.7 version
3	Spyder 4.2.3

8.2. Перечень информационно-справочных систем,используемых при осуществлении образовательного процесса по дисциплине

Перечень используемых информационно-справочных систем представлен в таблице 11.

Таблица 11– Перечень информационно-справочных систем

№ п/п	Наименование
1	Онлайн-книга «Справочник по науке о данных Python» https://jakevdp.github.io/PythonDataScienceHandbook/05.06-linear-regression.html
2	Датасеты по машинному обучению http://archive.ics.uci.edu/ml/datasets.php

9. Материально-техническая база

Состав материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине, представлен в таблице 12.

Таблица 12 – Состав материально-технической базы

№ п/п	Наименование составной части материально-технической базы	Номер аудитории (при необходимости)
1	Мультимедийная лекционная аудитория	Б.М. 21-21
2	Компьютерный класс	Б.М. 31-04

10. Оценочные средства для проведения промежуточной аттестации

10.1. Состав оценочных средств для проведения промежуточной аттестации обучающихся по дисциплине приведен в таблице 13.

Таблица 13 – Состав оценочных средств для проведения промежуточной аттестации

Вид промежуточной аттестации	Перечень оценочных средств
Дифференцированный зачёт	Список вопросов к дифф.зачёту; Примерный перечень вопросов для тестов.

10.2. В качестве критериев оценки уровня сформированности (освоения) компетенций обучающимися применяется 5-балльная шкала оценки сформированности компетенций, которая приведена в таблице 14. В течение семестра может использоваться 100-балльная шкала модульно-рейтинговой системы Университета, правила использования которой, установлены соответствующим локальным нормативным актом ГУАП.

Таблица 14 – Критерии оценки уровня сформированности компетенций

Оценка компетенции 5-балльная шкала	Характеристика сформированных компетенций
«отлично» «зачтено»	<ul style="list-style-type: none"> – обучающийся глубоко и всесторонне усвоил программный материал; – уверенно, логично, последовательно и грамотно его излагает; – опираясь на знания основной и дополнительной литературы, тесно привязывает усвоенные научные положения с практической деятельностью направления; – умело обосновывает и аргументирует выдвигаемые им идеи; – делает выводы и обобщения; – свободно владеет системой специализированных понятий.
«хорошо» «зачтено»	<ul style="list-style-type: none"> – обучающийся твердо усвоил программный материал, грамотно и по существу излагает его, опираясь на знания основной литературы; – не допускает существенных неточностей; – увязывает усвоенные знания с практической деятельностью направления; – аргументирует научные положения; – делает выводы и обобщения; – владеет системой специализированных понятий.
«удовлетворительно» «зачтено»	<ul style="list-style-type: none"> – обучающийся усвоил только основной программный материал, по существу излагает его, опираясь на знания только основной литературы; – допускает несущественные ошибки и неточности; – испытывает затруднения в практическом применении знаний направления; – слабо аргументирует научные положения;

Оценка компетенции	Характеристика сформированных компетенций
5-балльная шкала	
	<ul style="list-style-type: none"> – затрудняется в формулировании выводов и обобщений; – частично владеет системой специализированных понятий.
«неудовлетворительно» «не зачтено»	<ul style="list-style-type: none"> – обучающийся не усвоил значительной части программного материала; – допускает существенные ошибки и неточности при рассмотрении проблем в конкретном направлении; – испытывает трудности в практическом применении знаний; – не может аргументировать научные положения; – не формулирует выводов и обобщений.

10.3. Типовые контрольные задания или иные материалы.

Вопросы (задачи) для экзамена представлены в таблице 15.

Таблица 15 – Вопросы (задачи) для экзамена

№ п/п	Перечень вопросов (задач) для экзамена	Код индикатора
	Учебным планом не предусмотрено	

Вопросы для дифф. зачета представлены в таблице 16.

Таблица 16 – Вопросы для дифф. зачета

№ п/п	Перечень вопросов для дифф. зачета	Код индикатора
1	<ol style="list-style-type: none"> 1. Понятие «Искусственный интеллект». Технологии, которые принято относить к технологиям искусственного интеллекта. 2. Интеллектуальные задачи и их отличительные особенности. 3. Функции интеллекта человека. Характерные черты и свойства естественного интеллекта. Мышление. Задачи, решение которых связано с интеллектуальной деятельностью. 4. Поведенческая трактовка интеллекта и план имитации интеллекта А. Тьюринга. 5. Понятия «Данные» и «Знания». Обработка данных в условиях существенной априорной неполноты знаний об алгоритмах получения требуемого результата. 6. Области применения искусственного интеллекта. Свойства слабоформализованных задач. 7. Основные этапы истории развития систем искусственного интеллекта. Появление термина «Искусственный интеллект». Два направления развития искусственного интеллекта. 	УК-1.Д.1
2	<ol style="list-style-type: none"> 8. Понятие «Машинное обучение». Три составляющие машинного обучения. Машинное обучение и 	УК-1.Д.2

	<p>искусственный интеллект.</p> <p>9. Основные виды машинного обучения.</p> <p>10. Обучение с учителем. Основные решаемые задачи. Области применения.</p> <p>11. Обучение без учителя. Основные решаемые задачи. Области применения.</p> <p>12. Обучение с подкреплением. Основные решаемые задачи. Области применения.</p> <p>13. Основные этапы машинного обучения, их назначение и основные особенности.</p> <p>14. Предварительная обработка данных в системах машинного обучения. Назначение. Основные особенности.</p> <p>15. Этап «Обучение» в системах машинного обучения. Принципы формирования тренировочного набора данных. Использование перекрёстной проверки. Гиперпараметрическая оптимизация.</p> <p>16. Этап оценки обобщающей способности в системах машинного обучения. Понятия точности (precision) и верности (accuracy) функционирования моделей.</p>	
3	<p>17. Особенности выбора оптимальной модели машинного обучения.</p> <p>18. Типовая зависимость между сложностью модели машинного обучения и оценкой её эффективности.</p> <p>19. Типовая зависимость оценки эффективности моделей машинного обучения от размера обучающей последовательности.</p> <p>20. Строение и функции биологического нейрона.</p> <p>21. Персептрон Розенблата. Правило обучения персептрона.</p> <p>22. Адаптивный линейный нейрон. Обучение методом пакетного градиентного спуска.</p> <p>23. Адаптивный линейный нейрон. Обучение методом стохастического градиентного спуска. Мини-пакетное обучение.</p>	УК-1.Д.3
4	<p>24. Библиотека машинного обучения scikit-learn. Решаемые задачи. Реализуемые методы.</p> <p>25. Модель логистической регрессии и её обучение.</p> <p>26. Решение проблемы переобучения методом регуляризации. L2-регуляризация.</p> <p>27. Метод опорных векторов и его использование в задачах классификации. Классификация с мягким зазором.</p>	ОПК-2.Д.2

	28. Модель классификатора на основе дерева решений. 29. Классификатор на основе k ближайших соседей. 30. Наивный байесовский классификатор.	
--	---	--

Перечень тем для курсового проектирования/выполнения курсовой работы представлены в таблице 17.

Таблица 17 – Перечень тем для курсового проектирования/выполнения курсовой работы

№ п/п	Примерный перечень тем для курсового проектирования/выполнения курсовой работы
	Учебным планом не предусмотрено

Вопросы для проведения промежуточной аттестации в виде тестирования представлены в таблице 18.

Таблица 18 – Примерный перечень вопросов для тестов

№ п/п	Примерный перечень вопросов для тестов	Код индикатора
1	Через какую точку должна пройти ROC-кривая в случае идеального классификатора? А) (0,0) В) (0,1) С) (1,1) ANSWER: В	УК-1.Д.1 УК-1.Д.2 УК-1.Д.3 ОПК-2.Д.2
2	Какое дерево называется бинарным? А) Если при его построении используются только бинарные признаки В) Если оно имеет только два листа С) Если из любой его внутренней вершины выходит ровно два ребра ANSWER: С	УК-1.Д.1 УК-1.Д.2 УК-1.Д.3 ОПК-2.Д.2
3	Можно ли при построении бинарного дерева использовать небинарные признаки? А) Да В) Нет ANSWER: А	УК-1.Д.1 УК-1.Д.2 УК-1.Д.3 ОПК-2.Д.2
4	Что такое бутстрап? А) один из подходов к построению подвыборок В) метод обучения С) минимизируемый функционал ANSWER: А	УК-1.Д.1 УК-1.Д.2 УК-1.Д.3 ОПК-2.Д.2
5	Что такое бутстрапированная выборка? А) это некоторый сегмент обучающей выборки произвольной длины В) это выборка, полученная из обучающей выборки с	УК-1.Д.1 УК-1.Д.2 УК-1.Д.3 ОПК-2.Д.2

	<p>возможным повторением объектов</p> <p>С) это выборка, из которой удалены неизвестные значения признаков</p> <p>ANSWER: B</p>	
6	<p>Какой длины будет бутстрапированная выборка?</p> <p>А) больше длины исходной выборки В) меньше длины исходной выборки С) равной длине исходной выборки</p> <p>ANSWER: C</p>	<p>УК-1.Д.1 УК-1.Д.2 УК-1.Д.3 ОПК-2.Д.2</p>
7	<p>Какова зависимость корреляции между деревьями в случайном лесе от параметра q (размер случайного подмножества признаков)?</p> <p>А) Чем меньше q, тем больше корреляция В) Чем больше q, тем больше корреляция С) Данный параметр не влияет на корреляцию</p> <p>ANSWER: B</p>	<p>УК-1.Д.1 УК-1.Д.2 УК-1.Д.3 ОПК-2.Д.2</p>
8	<p>Градиентный спуск – это * при использовании в машинном обучении</p> <p>А) алгоритм оптимизации параметров модели В) алгоритм поиска минимального значения гиперпараметров С) алгоритм оптимизации архитектуры модели Д) алгоритм поиска максимального значения функции потерь</p> <p>ANSWER: A</p>	<p>УК-1.Д.1 УК-1.Д.2 УК-1.Д.3 ОПК-2.Д.2</p>
9	<p>Какой "девиз" у парадигмы ансамблирования?</p> <p>А) Много слабых моделей даёт результат лучше, чем одна сильная В) Много сильных моделей вместе даёт во всем лучший результат С) Много моделей даёт хуже результат, чем одна</p> <p>ANSWER: A</p>	<p>УК-1.Д.1 УК-1.Д.2 УК-1.Д.3 ОПК-2.Д.2</p>
10	<p>Оценка корреляции по Пирсону оказала значение 0 между переменными. Что это означает?</p> <p>А) Переменные не имеют линейной взаимосвязи В) Переменные имеют нелинейную взаимосвязь С) Переменные не имеют нелинейную взаимосвязь Д) Переменные имеют максимальную линейную взаимосвязь (так как значение оценки 1 - отсутствие взаимосвязи)</p> <p>ANSWER: A</p>	<p>УК-1.Д.1 УК-1.Д.2 УК-1.Д.3 ОПК-2.Д.2</p>
11	<p>Для чего применяется кросс-валидация:</p> <p>А) Поиск оптимальных гиперпараметров при малом количестве данных для обучения В) Поиск оптимальных гиперпараметров при большом количестве данных для обучения</p>	<p>УК-1.Д.1 УК-1.Д.2 УК-1.Д.3 ОПК-2.Д.2</p>

	<p>С) Поиск оптимальных гиперпараметров при отсутствии данных для обучения</p> <p>Д) Чтобы не использовать тестовую выборку для оценки метрик</p> <p>ANSWER: A</p>	
12	<p>В чем заключается метод случайных подпространств?</p> <p>А) Генерация случайного подмножества обучающей выборки для дальнейшего обучения базового алгоритма</p> <p>В) Получение случайного подмножества из исходного количества признаков для дальнейшего обучения базового алгоритма</p> <p>С) Генерация случайного подмножества базовых алгоритмов для получения агрегированного результата</p> <p>ANSWER: B</p>	<p>УК-1.Д.1</p> <p>УК-1.Д.2</p> <p>УК-1.Д.3</p> <p>ОПК-2.Д.2</p>
13	<p>Какая из метрик лучше подходит для поиска сильных (больших) ошибок модели в задаче регрессии:</p> <p>А) RMSE</p> <p>В) MAE</p> <p>С) R^2</p> <p>Д) F1</p> <p>ANSWER: A</p>	<p>УК-1.Д.1</p> <p>УК-1.Д.2</p> <p>УК-1.Д.3</p> <p>ОПК-2.Д.2</p>
14	<p>Допустим, мы предсказываем сорт цветков (один из трех: F, G, E) и оцениваем метрики класса F</p> <p>Какой из показателей классификации отражает количество правильно предсказанных ответов оцениваемого класс.</p> <p>А) True Positive</p> <p>В) False Positive</p> <p>С) True Negative</p> <p>Д) False Negative</p> <p>ANSWER: A</p>	<p>УК-1.Д.1</p> <p>УК-1.Д.2</p> <p>УК-1.Д.3</p> <p>ОПК-2.Д.2</p>
15	<p>Какая из метрик может принимать отрицательные значения?</p> <p>А) R^2</p> <p>В) MAE</p> <p>С) RMSE</p> <p>Д) Accuracy</p> <p>ANSWER: A</p>	<p>УК-1.Д.1</p> <p>УК-1.Д.2</p> <p>УК-1.Д.3</p> <p>ОПК-2.Д.2</p>
16	<p>При наличии сильного перекоса классов по целевой переменной (1000 записей класса А и 40 записей класса Б) какая из метрик может исказить оценку работы модели?</p> <p>А) accuracy</p> <p>В) recall</p> <p>С) precision</p> <p>Д) F1</p> <p>ANSWER: A</p>	<p>УК-1.Д.1</p> <p>УК-1.Д.2</p> <p>УК-1.Д.3</p> <p>ОПК-2.Д.2</p>

17	<p>Как бороться с эффектом переобучения?</p> <p>A) Сделать модель проще (менее комплексной) B) Взять меньше данных C) Сделать модель сложнее (более комплексной) D) Проверять метрики только по обучающей выборке</p> <p>ANSWER: A</p>	<p>УК-1.Д.1 УК-1.Д.2 УК-1.Д.3 ОПК-2.Д.2</p>
18	<p>Как связаны между собой плотность распределения $f(x)$ и функция распределения $F(x)$?</p> <p>A) Плотность распределения - это производная функции распределения B) Функция распределения - это производная плотности распределения C) Эти величины никак не связаны</p> <p>ANSWER: A</p>	<p>УК-1.Д.1 УК-1.Д.2 УК-1.Д.3 ОПК-2.Д.2</p>
19	<p>В чем заключается правило трёх сигм?</p> <p>A) Три среднеквадратических отклонения - это доверительный интервал для математического ожидания B) Вероятность того, что случайная величина примет значение, отклоняющееся от математического ожидания больше чем на три среднеквадратических отклонения пренебрежимо мала C) Три среднеквадратических отклонения - это максимально возможное значение случайной величины</p> <p>ANSWER: B</p>	<p>УК-1.Д.1 УК-1.Д.2 УК-1.Д.3 ОПК-2.Д.2</p>
20	<p>К какому типу признаков Вы бы отнесли рост человека (в см)?</p> <p>A) Бинарный B) Вещественный C) Категориальный D) Порядковый</p> <p>ANSWER: B</p>	<p>УК-1.Д.1 УК-1.Д.2 УК-1.Д.3 ОПК-2.Д.2</p>
21	<p>К какому типу признака Вы бы отнести цвет глаз человека?</p> <p>A) Бинарный B) Вещественный C) Категориальный D) Порядковый</p> <p>ANSWER: C</p>	<p>УК-1.Д.1 УК-1.Д.2 УК-1.Д.3 ОПК-2.Д.2</p>
22	<p>Какой из подходов регуляризации позволяет обнулить веса линейной модели, тем самым показать ненужность признака?</p> <p>A) Лассо B) L2 C) Ридж D) Тихонова</p> <p>ANSWER: A</p>	<p>УК-1.Д.1 УК-1.Д.2 УК-1.Д.3 ОПК-2.Д.2</p>

23	<p>Может ли решающее дерево восстанавливать нелинейные зависимости?</p> <p>А) Да В) Нет С) Только некоторые</p> <p>ANSWER: A</p>	УК-1.Д.1 УК-1.Д.2 УК-1.Д.3 ОПК-2.Д.2
24	<p>В чём заключается основное различие случайного леса и бэггинга на деревьях решений?</p> <p>А) В случайном лесу для каждого разбиения признаки выбираются из некоторого случайного подмножества признаков В) В случайном лесу обучение идёт на бутстрапированной выборке С) Эти два метода ничем не отличаются</p> <p>ANSWER: A</p>	УК-1.Д.1 УК-1.Д.2 УК-1.Д.3 ОПК-2.Д.2
25	<p>Какие простые модели использует метод случайных лесов в качестве основы:</p> <p>А) Решающие деревья В) Линейная регрессия С) Логистическая регрессия D) Модель опорных векторов</p> <p>ANSWER: A</p>	УК-1.Д.1 УК-1.Д.2 УК-1.Д.3 ОПК-2.Д.2
26	<p>В чём особенность стохастического градиентного спуска?</p> <p>А) Градиент функции качества вычисляется только на одном случайно выбранном объекте обучающей выборки В) Градиент функции качества вычисляется со случайно выбранным шагом С) Градиент вычисляется для случайной функции качества</p> <p>ANSWER: A</p>	УК-1.Д.1 УК-1.Д.2 УК-1.Д.3 ОПК-2.Д.2
27	<p>Какая из задач предполагает восстановление моделью зависимости в данных путем приближения предсказаний модели к значениям в обучающих данных:</p> <p>А) регрессия В) классификация С) кластеризация D) прокрастинация</p> <p>ANSWER: A</p>	УК-1.Д.1 УК-1.Д.2 УК-1.Д.3 ОПК-2.Д.2
28	<p>Какая из задач предполагает поиск наилучшего разделения данных в пространстве признаков?</p> <p>А) классификация В) регрессия С) кластеризация D) валидация</p>	УК-1.Д.1 УК-1.Д.2 УК-1.Д.3 ОПК-2.Д.2

	ANSWER: A	
29	<p>Признак Grade имеет следующие уникальные значения: 1, 2, 3, 4, 5 и отражает оценку за тест студентов. Определите тип признака:</p> <p>A) Номинальный категориальный B) Последовательный категориальный C) Численный D) Дихотомический E) Нет правильного ответа</p> <p>ANSWER: B</p>	УК-1.Д.1 УК-1.Д.2 УК-1.Д.3 ОПК-2.Д.2
30	<p>В каком из случаев имеет смысл сразу исключить колонку данных:</p> <p>A) Вся колонка имеет одно и то же значение B) В колонке каждое значение - уникальное (вещественные числа) C) Колонка имеет два уникальных значения D) Колонка содержит строки</p> <p>ANSWER: A</p>	УК-1.Д.1 УК-1.Д.2 УК-1.Д.3 ОПК-2.Д.2

Перечень тем контрольных работ по дисциплине обучающихся заочной формы обучения, представлены в таблице 19.

Таблица 19 – Перечень контрольных работ

№ п/п	Перечень контрольных работ
	Не предусмотрено

10.4. Методические материалы, определяющие процедуры оценивания индикаторов, характеризующих этапы формирования компетенций, содержатся в локальных нормативных актах ГУАП, регламентирующих порядок и процедуру проведения текущего контроля успеваемости и промежуточной аттестации обучающихся ГУАП.

11. Методические указания для обучающихся по освоению дисциплины

11.1. Методические указания для обучающихся по освоению лекционного материала

Основное назначение лекционного материала – логически стройное, системное, глубокое и ясное изложение учебного материала. Назначение современной лекции в рамках дисциплины не в том, чтобы получить всю информацию по теме, а в освоении фундаментальных проблем дисциплины, методов научного познания, новейших достижений научной мысли. В учебном процессе лекция выполняет методологическую, организационную и информационную функции. Лекция раскрывает понятийный аппарат конкретной области знания, её проблемы, дает цельное представление о дисциплине, показывает взаимосвязь с другими дисциплинами.

Планируемые результаты при освоении обучающимися лекционного материала:

- получение современных, целостных, взаимосвязанных знаний, уровень которых определяется целевой установкой к каждой конкретной теме;
- получение опыта творческой работы совместно с преподавателем;
- развитие профессионально-деловых качеств, любви к предмету и самостоятельного творческого мышления.

- появление необходимого интереса, необходимого для самостоятельной работы;
- получение знаний о современном уровне развития науки и техники и о прогнозе их развития на ближайшие годы;
- научиться методически обрабатывать материал (выделять главные мысли и положения, приходиться к конкретным выводам, повторять их в различных формулировках);
- получение точного понимания всех необходимых терминов и понятий.

Лекционный материал может сопровождаться демонстрацией слайдов и использованием раздаточного материала при проведении коротких дискуссий об особенностях применения отдельных тематик по дисциплине.

Структура предоставления лекционного материала:

Лекция состоит из трёх основных частей: вступительной, основной и заключительной.

Вступительная часть определяет название темы, план и цель лекции. Она призвана заинтересовать и настроить аудиторию. В этой части лекции излагается актуальность, основная идея, связь данной лекции с предыдущими занятиями, ее основные вопросы. Введение должно быть кратким и целенаправленным.

В основной части лекции реализуется научное содержание темы, все главные узловые вопросы, проводится вся система доказательств с использованием наиболее целесообразных методических приёмов. Каждый учебный вопрос заканчивается краткими выводами, логически подводящими студентов к следующему вопросу лекции.

Заключительная часть имеет целью обобщать в кратких формулировках основные идеи лекции, логически завершая её как целостное творение.

Отдельные виды лекций могут иметь свои особенности как по содержанию, так и по структуре.

11.2. Методические указания для обучающихся по прохождению практических занятий

Практическое занятие является одной из основных форм организации учебного процесса, заключающаяся в выполнении обучающимися под руководством преподавателя комплекса учебных заданий с целью усвоения научно-теоретических основ учебной дисциплины, приобретения умений и навыков, опыта творческой деятельности.

Целью практического занятия для обучающегося является привитие обучающимся умений и навыков практической деятельности по изучаемой дисциплине.

Планируемые результаты при освоении обучающимся практических занятий:

- закрепление, углубление, расширение и детализация знаний при решении конкретных задач;
- развитие познавательных способностей, самостоятельности мышления, творческой активности;
- овладение новыми методами и методиками изучения конкретной учебной дисциплины;
- выработка способности логического осмысления полученных знаний для выполнения заданий;
- обеспечение рационального сочетания коллективной и индивидуальной форм обучения.

Требования к проведению практических занятий

Практические занятия проводятся в компьютерном классе с использованием пакетов программ разработки и отладки программного обеспечения дистрибутива Anaconda.

Anaconda – Дистрибутив Python, предназначенный для крупномасштабной обработки данных, прогнозной аналитики и научных вычислений от компании Continuum

Analytics. Это бесплатный, включая коммерческое использование, и готовый к использованию в среде предприятия дистрибутив Python, который объединяет все ключевые библиотеки, необходимые для работы в области науки о данных, математики и разработки. Anaconda уже включает NumPy, SciPy, matplotlib, pandas, IPython, Jupyter Notebook и scikit-learn.

Практические занятия предполагают работу по индивидуальному заданию, связанному с разработкой и отладкой робототехнических систем обработки информации.

По результатам каждого практического задания должен быть подготовлен отчёт, содержащий необходимые теоретические сведения, листинги m-программ, построенные в процессе выполнения работы графические зависимости и т.п. Каждый отчёт должен содержать выводы по проделанной работе и список используемых дополнительных источников.

Каждое практическое занятие завершается собеседованием с преподавателем по представленному отчёту с выставлением оценки.

11.3. Методические указания для обучающихся по прохождению самостоятельной работы

В ходе выполнения самостоятельной работы, обучающийся выполняет работу по заданию и при методическом руководстве преподавателя, но без его непосредственного участия.

В процессе выполнения самостоятельной работы, у обучающегося формируется целесообразное планирование рабочего времени, которое позволяет им развивать умения и навыки в усвоении и систематизации приобретаемых знаний, обеспечивает высокий уровень успеваемости в период обучения, помогает получить навыки повышения профессионального уровня.

Методическими материалами, направляющими самостоятельную работу обучающихся являются:

- учебно-методический материал по дисциплине.

11.4. Методические указания для обучающихся по прохождению текущего контроля успеваемости

Текущий контроль успеваемости предусматривает контроль качества знаний обучающихся, осуществляемого в течение семестра с целью оценивания хода освоения дисциплины.

Контроль качества знаний проводится в форме индивидуального собеседования по материалу отдельных разделов дисциплины, а также проверки отчётов о выполнении практических заданий.

Результаты текущего контроля могут учитываться при проведении промежуточной аттестации. Положительный результат текущего контроля успеваемости дает студенту дополнительный балл при проведении промежуточной аттестации. Результаты текущего контроля могут учитываться при проведении промежуточной аттестации.

Система оценок при проведении промежуточной аттестации осуществляется в соответствии с требованиями Положений «О текущем контроле успеваемости и промежуточной аттестации студентов ГУАП, обучающихся по программы высшего образования» и «О модульно-рейтинговой системе оценки качества учебной работы студентов в ГУАП».

11.5. Методические указания для обучающихся по прохождению промежуточной аттестации

Промежуточная аттестация обучающихся предусматривает оценивание промежуточных и окончательных результатов обучения по дисциплине. Она проводится в форме дифференцированного зачёта.

Дифференцированный зачет – это форма оценки знаний, полученных обучающимся при изучении дисциплины, при выполнении курсовых проектов, курсовых работ, научно-исследовательских работ и прохождении практик с аттестационной оценкой «отлично», «хорошо», «удовлетворительно», «неудовлетворительно».

Система оценок при проведении промежуточной аттестации осуществляется в соответствии с требованиями Положений «О текущем контроле успеваемости и промежуточной аттестации студентов ГУАП, обучающихся по программам высшего образования» и «О модульно-рейтинговой системе оценки качества учебной работы студентов в ГУАП».

Промежуточная аттестация оценивается по результатам текущего контроля успеваемости. Список вопросов (таблица 16) к промежуточной аттестации утверждается кафедрой и выдается студентам для ознакомления. В случае, если студент по уважительной причине не выполнил требования текущего контроля, ему предоставляется возможность сдать задолженности по пропущенным темам. Форма проведения промежуточной аттестации – письменная.

Лист внесения изменений в рабочую программу дисциплины

Дата внесения изменений и дополнений. Подпись внесшего изменения	Содержание изменений и дополнений	Дата и № протокола заседания кафедры	Подпись зав. кафедрой