

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ  
ПОВОЛЖСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНОЛОГИЧЕСКИЙ УНИВЕРСИТЕТ



УТВЕРЖДАЮ  
Директор ИММ

УТВЕРЖДАЮ /Н.П. Сютлов/  
(Ф.И.О. декана (директора института))

02.02.2022 г.

**РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)**

**Б.1.2.3 Анализ больших данных**

*(код и наименование дисциплины по учебному плану)*

Направление подготовки  
(специальность)

35.03.06 Агроинженерия

Квалификация выпускника

Бакалавр

*(бакалавр/магистр/специалист)*

Направленность

Искусственный интеллект в агроинженерии

Курс 3  
Семестр 5

**Распределение учебного времени**

Трудоемкость по учебному плану	180 / 5	часов/зачетных единиц
Лекции	16	часов
Лабораторные работы	-	часов
Практические занятия	64	часов
Иная контактная работа	-	часов
Всего контактной работы (без учета экз.)	80	часов
Контактная работа по экзамену	6	часов
Курсовой проект (работа)	-	семестр
Самостоятельная работа обучающихся (без учета экз.)	64	часов
Самостоятельная работа по подготовке к экзамену	30	часов
Экзамен	5	семестр
Зачет	-	семестр
БРК, ДЗ	-	семестр

                      
(год)

Программа составлена в соответствии с требованиями ФГОС ВО направления подготовки (специальности) 35.03.06 Агроинженерия

Программу составили:

заведующий кафедрой с ученой степенью кандидата наук	ЭМиО	СОГЛАСОВАНО	Д.В. Костромин
(должность)	(кафедра)		(И.О. Фамилия)

РАССМОТРЕНА и ОДОБРЕНА на заседании кафедры, за которой закреплена дисциплина  
Кафедра информатики

		(наименование кафедры)	
11.10.2021	протокол №	3	
(дата)			
Заведующий кафедрой	СОГЛАСОВАНО	А.В. Кревецкий	
		(И.О. Фамилия)	

Рабочая программа СОГЛАСОВАНА с факультетом (институтом), выпускающей(ими)  
кафедрой(ами).  
СООТВЕТСТВУЕТ действующей ОП.

Заведующий кафедрой	СОГЛАСОВАНО	Д.В. Костромин
		(И.О. Фамилия)

Председатель методической комиссии факультета (института), в который входит  
выпускающая кафедра

СОГЛАСОВАНО	А.А. Медяков
	(И.О. Фамилия)

Эксперт(ы): Усков Юрий Викторович, Генеральный директор ООО "Ричмедиа"

Рабочая программа проверена и зарегистрирована в УМЦ 07.02.2022 г.  
Специалист учебно-методического центра СОГЛАСОВАНО /Т.А. Смирнова/

## Раздел 1. ЦЕЛЬ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Целью освоения дисциплины является достижение планируемых результатов обучения, соответствующих установленным в ОПОП индикаторам достижения компетенций:

Код и наименование компетенции	Код и наименование индикатора достижения компетенции	Результаты обучения
1. ПК-2 Способен классифицировать и идентифицировать задачи искусственного интеллекта, выбирать адекватные методы и инструментальные средства решения задач искусственного интеллекта	ПК-2.1 Классифицирует и идентифицирует задачи систем искусственного интеллекта в зависимости от особенностей проблемной и предметной областей	<b>знания:</b> методологические подходы извлечения знаний из данных и основы применения соответствующих инструментальных средств <b>умения:</b> умеет выбрать соответствующие инструментальные средства и платформы <b>навыки:</b> имеет навык выбора инструментальных средств и платформ, основанных на извлечении знаний, для решения поставленных задач
2. ПК-5 Способен осуществлять сбор и подготовку данных для систем искусственного интеллекта	ПК-5.1 Осуществляет поиск данных в открытых источниках, специализированных библиотеках и репозиториях.	<b>знания:</b> принципы и методы машинного обучения, типы и классы задач машинного обучения <b>умения:</b> сопоставлять задачам предметной области классы задач машинного обучения <b>навыки:</b> постановки задач по разработке или совершенствованию методов и алгоритмов для решения комплекса задач
	ПК-5.2 Выполняет подготовку и разметку структурированных и неструктурированных данных для машинного обучения.	<b>знания:</b> принципы и методы машинного обучения, типы и классы задач машинного обучения <b>умения:</b> сопоставлять задачам предметной области классы задач машинного обучения <b>навыки:</b> постановки задач по разработке или совершенствованию методов и алгоритмов для решения комплекса задач
3. ПК-6 Способен выполнять анализ больших данных	ПК-6.1 Использует знания о вариантах использования больших данных, определениях, словарях и эталонной архитектуре больших данных для эффективного извлечения, хранения и подготовки больших данных.	<b>знания:</b> функциональные возможности современных инструментальных средств и систем программирования в области создания моделей и методов машинного обучения <b>умения:</b> проводить сравнительный анализ и осуществлять выбор инструментальных средств для решения задач машинного обучения, в частности для решения задач распознавания объектов по звуку и по изображениям <b>навыки:</b> разработки требований к

## Раздел 2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОПОП

Дисциплина относится к части, формируемой участниками образовательных отношений ОПОП.

Дисциплина является обязательной

Изучаемая дисциплина является основой для продолжения формирования указанных компетенций в следующих дисциплинах: Основы программирования систем искусственного интеллекта на Python (ПК-5), Интеллектуальные системы управления технологическими процессами в АПК (ПК-2); государственной итоговой аттестации в форме: Выполнение и защита выпускной квалификационной работы (ПК-6)

## Раздел 3. ОПИСАНИЕ ОБРАЗОВАТЕЛЬНЫХ ТЕХНОЛОГИЙ

Для формирования заявленных компетенций используются методологические технологии, реализующие деятельностный, личностно-ориентированный, практико-ориентированный подходы.

Основными стратегическими технологиями являются: имитационное моделирование, исследовательские, лекционные занятия, практические и лабораторные занятия, процедуры самообучения

На достижение конкретных целей обучения направлены применяемые тактические технологии: задания, классическая лекция, мини-проекты

## Раздел 4. СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

### 5 семестр

Виды и темы занятий	Количество часов	Формируемые компетенции
<b>Введение в анализ данных на Python</b>	<b>34</b>	ПК-2, ПК-5, ПК-6
Лекция. Пакеты Python для науки о данных	1	
Практическое занятие. Предварительная обработка данных больших массивов данных (NumPy, Pandas)	8	
Лекция. Меры центральной тенденции	1	
Практическое занятие. Вычисление и сравнение статистических характеристик различных массивов данных, их визуализация	8	
Задания для самостоятельной работы, в том числе выполнение Пакеты Python для науки о данных Предварительная обработка Визуализация данных Меры центральной тенденции	16	
<b>Временные ряды и регрессионный анализ</b>	<b>33</b>	ПК-2, ПК-5, ПК-6
Лекция. Временные ряды и задачи по их обработке	1	
Лекция. Интерполяция и экстраполяция функций	2	
Практическое занятие. Интерполяция функции сплайнами	6	
Практическое занятие. Метод опорных векторов в решении задач регрессии	8	
Практическое занятие. Корреляционный анализ	8	

Задания для самостоятельной работы, в том числе выполнение Понятие временного ряда Интерполяция функций Экстраполяция функций (регрессионные модели) Авторегрессионные модели Корреляция	8	ПК-2, ПК-5, ПК-6
<b>Кластеризация и классификация</b>	<b>32</b>	
Лекция. Методы решения задачи классификации	1	
Лекция. Классификация и кластеризация с подкреплением	1	
Практическое занятие. Методы самообучения в задачах классификации	8	
Практическое занятие. Кластеризация методом k средних	8	ПК-2, ПК-5, ПК-6
Задания для самостоятельной работы, в том числе выполнение Методы решения задачи классификации Классификация с подкреплением Кластеризация с подкреплением	14	
<b>Интеллектуальные технологии</b>	<b>21</b>	
Лекция. Элементы теории автоматов	1	
Лекция. Машинное обучение. Нейронные сети	2	
Лекция. Экспертные системы	1	ПК-2, ПК-5, ПК-6
Лекция. Введение в машинное обучение. Построение нейронной сети на PyTorch	3	
Задания для самостоятельной работы, в том числе выполнение РГР Элементы теории автоматов Сети Петри Машинное обучение Нейронные сети Экспертные системы Деревья решений Наивный байесовский классификатор	14	
<b>Оценка качества моделей</b>	<b>24</b>	
Лекция. Проверка адекватности в задачах регрессии	2	ПК-2, ПК-5, ПК-6
Практическое занятие. Проверка моделей (гипотез) по критериям согласия	10	
Задания для самостоятельной работы, в том числе выполнение Критерии согласия в задачах регрессии Законы распределения Метрики в задачах классификации	12	
Иная контактная работа: выполнение контрольной работы	0	
Подготовка к экзамену	30	
Проведение экзамена	6	

## Раздел 5. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ОБУЧАЮЩИХСЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ

Изучение дисциплины рекомендуется начать с ознакомления с рабочей программой, ее структурой и содержанием разделов. Учебный материал структурирован, изучение

дисциплины осуществляется в тематической последовательности. **Занятия лекционного типа** дают систематизированные знания по дисциплине, концентрируют внимание на наиболее сложных и важных вопросах. Во время лекционных занятий рекомендуется вести конспектирование учебного материала; обращать внимание на формулировки и категории, раскрывающие суть проблемы, явления или процесса; зафиксировать выводы и практические рекомендации. Подготовка к занятиям **семинарского типа** включает ознакомление с планом практического занятия; работу с конспектом лекций, выполнение домашнего задания, работу с учебной и учебно-методической литературой, научными изданиями и электронными образовательными ресурсами, рекомендованными рабочей программой дисциплины.

Содержание **самостоятельной работы** определяется рабочей программой дисциплины, оценочными и методическими материалами, заданиями и указаниями преподавателя. Самостоятельная работа может осуществляться в аудиторной и внеаудиторной формах. Эффективным средством осуществления самостоятельной работы является электронная информационно-образовательная среда университета, которая обеспечивает доступ к образовательной программе, рабочей программе дисциплины, к электронным библиотечным системам, профессиональным базам данных и информационным справочным системам.

Изучение дисциплины включает выполнение расчётно-графической работы, контрольной работы, практической работы и т.д. Периодичность проведения, формы текущего контроля успеваемости, система оценивания хода освоения дисциплин представлены в рабочей программе.

## Раздел 6. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ И УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

### 6.1. Учебно-методическое обеспечение

№№ п/п	Список используемой литературы	Количество экземпляров печатных изданий, имеющих в библиотеке, или электронный адрес издания (ресурса) в сети Интернет
<b>УЧЕБНЫЕ, УЧЕБНО-МЕТОДИЧЕСКИЕ И НАУЧНЫЕ ИЗДАНИЯ</b>		
1.	Есипов, Борис Алексеевич. Методы исследования операций [Текст] : учеб. пособие / Б. А. Есипов. Санкт-Петербург [и др.]: Лань, 2010. - 253 с. ISBN 978-5-8114-0917-4. Экземпляры: всего 6.	6
2.	Рутковская, Данута. Нейронные сети, генетические алгоритмы и нечеткие системы [Текст] / Д. Рутковская, М. Пилиньский, Л. Рутковский ; пер. с пол. И. Д. Рудинского. М.: Горячая линия - Телеком, 2007. - 383 с. ISBN 5-93517-103-1. Экземпляры: всего 4.	4
3.	Хайкин, Саймон. Нейронные сети [Текст] : Полный курс / С. Хайкин ; [пер. с англ. Н. Н. Куссуль, А. Ю. Шелестова ; под ред. Н. Н. Куссуль]. 2-е изд. Москва [и др.]: Вильямс, 2006. - 1103 с. ISBN 5-8459-0890-6.	8
4.	Барский, Аркадий Бенционович. Нейронные сети [Текст] : распознавание, управление, принятие решений / А. Б. Барский. М.: Финансы и статистика, 2004. - 175 с. ISBN 5-279-02757-X. Экземпляры: всего 9.	9
5.	Чубукова, И. А. Data Mining [Электронный ресурс] /	

	Чубукова И. А. 2-е изд. Москва: ИНТУИТ, 2016. - 470 с. ISBN 978-5-94774-819-2.	<a href="https://e.lanbook.com/book/100582">https://e.lanbook.com/book/100582</a>
6.	Макшанов, А. В. Большие данные. Big Data [Электронный ресурс] / Макшанов А. В., Журавлев А. Е., Тындыкарь Л. Н. 2-е изд., стер. Санкт-Петербург: Лань, 2022. - 188 с. ISBN 978-5-8114-9690-7.	<a href="https://e.lanbook.com/book/198599">https://e.lanbook.com/book/198599</a>
7.	Макшанов, А. В. Технологии интеллектуального анализа данных [Электронный ресурс] : учебное пособие / А. В. Макшанов, А. Е. Журавлев. 2-е изд., стер. Санкт-Петербург: Лань, 2022. - 212 с. ISBN 978-5-8114-4493-9.	<a href="https://e.lanbook.com/book/206711">https://e.lanbook.com/book/206711</a>

## 6.2. Материально-техническая база и программное обеспечение

№№ п/п	Аудитории для проведения учебных занятий, самостоятельной работы и проведения государственной итоговой аттестации	Перечень основного оборудования	Программное обеспечение
1.	119 (II)	ПК ICL RAY S902.1, клавиат., мышь, патч корд 3м, монитор ViewSonic 21,5" VA2248-LED (15), Стойка компьютерная (15), Комплект учебной мебели (1)	Microsoft Windows Enterprise, Справочная правовая система "Консультант Плюс", Microsoft Office Standard, Агент Dr.Web, Комплект ГАРАНТ-Мастер, Microsoft Access, Microsoft Visio Professional, Microsoft Project Professional, Microsoft Visual Studio Enterprise, Комплект ПО для решения основных пользовательских задач

## Раздел 7. ФОРМЫ КОНТРОЛЯ ОСВОЕНИЯ ДИСЦИПЛИНЫ/ ФОНД ОЦЕНОЧНЫХ СРЕДСТВ

Критерии оценивания индикаторов достижения компетенций направлены на:

- усвоение теоретического материала (объем знаний, глубина усвоения), предусмотренного рабочей программой;
- умение излагать материал (четкость, грамотность изложения материала, точность и полнота воспроизведения учебного материала);
- умение применять теоретические знания при решении практических заданий.

Шкала оценивания представлена ниже.

Уровень сформированности элементов компетенции	Критерии оценивания	Шкала оценивания
Пороговый уровень	Обучающийся имеет знания основного материала, проявляет умение логично его излагать, но может допускать неточности в изложении материала, недостаточно правильные формулировки, испытывает затруднения в выполнении практических заданий.	удовлетворительно

Продвинутый уровень	Обучающийся твердо знает программный материал, излагает его грамотно и по существу, не допускает существенных неточностей в ответе на вопрос, правильно применяет теоретические положения при решении практических вопросов и задач, владеет необходимыми навыками и приемами их выполнения	хорошо
Высокий уровень	Обучающийся глубоко и прочно усвоил программный материал, грамотно и логически стройно его излагает, дает исчерпывающие ответы на поставленные вопросы. В ответе тесно увязывается теория с практикой, при этом обучающийся не затрудняется с ответом при видоизменении задания, свободно справляется с задачами, вопросами и другими видами применения знаний, показывает знакомство с монографической литературой, периодическими изданиями, правильно обосновывает принятые решения, свободно владеет разносторонними навыками, приемами выполнения практических работ	отлично

### 7.1. Текущий контроль успеваемости

Текущий контроль успеваемости обеспечивает оценивание хода освоения дисциплины (модуля) и производится с применением технологии рейтингового контроля в соответствии с технологической картой дисциплины. Порядок составления технологической карты и алгоритм проведения процедуры оценивания видов деятельности обучающихся, направленных на освоение знаний, умений, навыков и/или опыта деятельности, по накопительной системе в баллах устанавливается положением о системе РИТМ в ФГБОУ ВО «ПГТУ»

### 7.2. Промежуточная аттестация обучающихся

Промежуточная аттестация обучающихся направлена на оценивание результатов обучения по дисциплине (модулю) и проводится с использованием фондов оценочных средств.

Примеры типовых контрольных заданий из базы фонда оценочных средств по образовательной программе.

#### 1. Машинное обучение - это ...

- научное направление, задачей которого является создание интеллектуальных систем, лежит на стыке информатики, статистики и анализа данных, а также занимается вопросами, связанными с философией и этичностью использования интеллектуальных систем
- объединение ряда научных областей, занимающихся построением систем анализа и обработки данных
- математическая область, связанная с построением предсказательных алгоритмов (как правило представленных статистическими моделями) на основе данных
- прикладное направление, развивающее методы построения баз знаний и правил, явным образом описывающих знания экспертов

#### 2. Какие задачи из ниже перечисленных относятся к задачам классификации?

- определение наиболее целесообразного способа лечения
- определение длительности и исхода заболевания



- задачи поискового вывода

- оценивание кредитоспособности заёмщика

3. Определения дохода клиента банка является задачей ...

- регрессии

- кластеризации

- классификации

4. В чём заключается задача кластеризации?

- задача машинного обучения, в которой метки объектов принимают ограниченное число значений, например, город проживания, пол клиента

- задача машинного обучения, в которой метки объектов принимают любое численное значение, например, стоимость квартиры, сумма кредита

- задача машинного обучения, заключающаяся в объединении похожих объектов в однородные группы

5. Что называют обучением с подкреплением?

- обучение с контролирующими параметрами

- обучение с большим количеством учителей

- обучение основанное на собственном опыте

6. У машинного обучения есть ряд задач. Как называется та, что направлена на предсказание значения той или иной непрерывной числовой величины для входных данных?

- Переобучение

- Классификация

- Регрессия

- Кластеризация

7. Нейросети хорошо проявляют себя не только в распознавании, но и в генерации изображений. Но кое с чем у них все-таки возникают проблемы. С чем именно?

- Цвет

- Текстуры

- Формы

- Глубина, количество пикселей

8. Какой из видов машинного обучения основывается на взаимодействии обучаемой системы со средой?

- Обучение с подкреплением

- Глубинное обучение

- Обучение с учителем

- Обучение без учителя

9. Когда говорят о нейронных сетях и машинном обучении, часто упоминают закон Мура. В чем его суть?

- 20% усилий дают 80% результата, а остальные 80% усилий — лишь 20% результата
- Если все слова языка или длинного текста упорядочить по убыванию частоты их использования, то частота  $n$ -го слова в таком списке окажется приблизительно обратно пропорциональной его порядковому номеру  $n$
- Не следует множить сущее без необходимости
- Каждое следующее поколение компьютеров работает в 2,5 раза быстрее

10. Какой тип искусственной нейронной сети представлен на картинке?

- Нейронная сеть Джордана
- Рекуррентная нейронная сеть
- Простая нейронная сеть
- Сверточная нейронная сеть

### Перечень вопросов для проведения промежуточной аттестации

1. Найдите значения коэффициентов  $c_1$  и  $c_2$  для функции  $y=c_1+c_2x$ , если известно, что функция экстраполирует данные, приведенные в таблице.
2. Постройте экстраполяционную модель с помощью кривых Гомперца и Перла, описывающую данные Росстата о количестве автомобилей на тысячу человек при условии, что значение должно приближаться к значению 1000.
3. С помощью теории автоматов опишите работу вендинговой машины (автоматапродавца) по продаже двух товаров по цене 5 и 10 рублей, при условии, что машина принимает монеты номиналов 1, 2, 5 и 10 рублей. Учтите, что выдача товара должна происходить при нажатии кнопки выбора соответствующего товара. Усложните автомат таким образом, чтобы он выдавал сдачу.
4. С помощью нейронной сети опишите работу вендинговой машины (автоматапродавца) по продаже двух товаров по цене 5 и 10 рублей, при условии, что машина принимает монеты номиналов 1, 2, 5 и 10 рублей. Учтите, что выдача товара должна происходить при нажатии кнопки выбора соответствующего товара. Усложните автомат таким образом, чтобы он выдавал сдачу.
5. На рисунке приведены объекты, относящиеся к разным классам (треугольники и круги), определяемые на основе значений пары параметров  $y$  и  $x$ . Запишите функцию  $y = f(x)$ , которая позволит разделять приведенные данные на классы.
6. Постройте матрицу признаков для экспертной системы матричного типа, которая позволит идентифицировать несколько простудных заболеваний (например, ОРЗ, грипп, ангина) по таким признакам, как температура тела, головная боль, кашель, ломота в суставах, сонливость, насморк, боль в горле.
7. Проверьте модели, полученные в результате выполнения заданий, с помощью критериев согласия Фишера, Пирсона, Стьюдента и Колмогорова — Смирнова. Сравните полученные результаты, сделайте выводы о применимости критериев.
8. Проверьте модели, полученные в результате выполнения заданий, с использованием Accurasy, F-меры и логистической функции потерь. Сравните результаты, сделайте выводы о применимости

критериев оценки моделей.