

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ  
ПОВОЛЖСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНОЛОГИЧЕСКИЙ УНИВЕРСИТЕТ



УТВЕРЖДАЮ  
Декан ФУП

УТВЕРЖДАЮ /О.М. Репина/  
(Ф.И.О. декана (директора института))

11.03.2024 г.

**РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)**

М.1.2.1 Машинное обучение в управлении организацией

(код и наименование дисциплины по учебному плану)

Направление подготовки  
(специальность)

27.04.02 Управление качеством

Квалификация выпускника

Магистр

(бакалавр/магистр/специалист)

Программа магистратуры

Искусственный интеллект в менеджменте качества

Курс 2  
Семестр 3

**Распределение учебного времени**

Трудоемкость по учебному плану	180 / 5	часов/зачетных единиц
Лекции	12	часов
Лабораторные работы	-	часов
Практические занятия	24	часов
Иная контактная работа	-	часов
Всего контактной работы (без учета экз.)	36	часов
Контактная работа по экзамену	6	часов
Курсовой проект (работа)	-	семестр
Самостоятельная работа обучающихся (без учета экз.)	108	часов
Самостоятельная работа по подготовке к экзамену	30	часов
Экзамен	3	семестр
Зачет	-	семестр
БРК, ДЗ	-	семестр

(год)

Программа составлена в соответствии с требованиями ФГОС ВО направления подготовки (специальности) 27.04.02 Управление качеством

Программу составили:

старший преподаватель	ПиП ЭВС	СОГЛАСОВАНО	Н.В. Скобелева
(должность)	(кафедра)		(И.О. Фамилия)

РАССМОТРЕНА и ОДОБРЕНА на заседании кафедры, за которой закреплена дисциплина  
Кафедра проектирования и производства электронно-вычислительных средств

(наименование кафедры)		
05.02.2024	протокол №	9
(дата)		

Заведующий кафедрой	СОГЛАСОВАНО	Т.С. Буканова
		(И.О. Фамилия)

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

СООТВЕТСТВУЕТ действующей ОП.

Заведующий кафедрой	СОГЛАСОВАНО	Т.В. Ялялиева
		(И.О. Фамилия)

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

СОГЛАСОВАНО	С.В. Краснова
	(И.О. Фамилия)

Эксперт(ы): Поздеев Сергей Валерьевич, Директор АНО "РЦК в сфере производительности труда в РМЭ"

Рабочая программа проверена и зарегистрирована в УМЦ 12.03.2024 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-3 Способен выбирать и участвовать в проведении экспериментальной проверки работоспособности программных платформ систем искусственного интеллекта по обеспечению требуемых критериев эффективности и качества функционирования	ПК-3.1. Выбирает программные платформы систем искусственного интеллекта	<b>знания:</b> основные критерии эффективности и качества функционирования системы искусственного интеллекта: точность, релевантность, достоверность, целостность, быстрота решения задач, надежность, защищенность функционирования <b>умения:</b> выбирать и применять программные платформы систем искусственного интеллекта с учетом основных критериев эффективности и качества функционирования <b>навыки:</b> выбора и применения программных платформ искусственного интеллекта
	ПК-3.2. Участвует в проведении экспериментальной проверки работоспособности систем искусственного интеллекта	<b>знания:</b> методы постановки задач, проведения анализа тестовых и экспериментальных испытаний работоспособности систем искусственного интеллекта <b>умения:</b> ставить задачи и участвовать в проведении тестовых и экспериментальных испытаний работоспособности систем искусственного интеллекта, анализировать результаты и вносить изменения <b>навыки:</b> проведения тестовых и экспериментальных испытаний работоспособности систем искусственного интеллекта
2. ПК-4 Способен управлять проектами по созданию, поддержке и использованию систем искусственного интеллекта со стороны заказчика	ПК-4.1. Организует работы по управлению проектами создания, внедрения и использования систем искусственного интеллекта со стороны заказчика	<b>знания:</b> методы и средства управления проектами создания, внедрения и использования систем искусственного интеллекта со стороны заказчика с учетом рисков, возникающих во внутренней и внешней среде <b>умения:</b> применять методы и средства управления проектами создания, внедрения и использования систем искусственного интеллекта со стороны заказчика с учетом рисков, возникающих во внешней и внутренней среде <b>навыки:</b> использования методов и средств управления проектами создания, внедрения и использования систем искусственного интеллекта со стороны заказчика с учетом рисков, возникающих во внешней и внутренней среде

	ПК-4.2. Организует и руководит коллективной работой по созданию, внедрению и использованию систем искусственного интеллекта со стороны заказчика	<b>знания:</b> методы и средства взаимодействия с инженерами по знаниям, разработчиками, ключевыми пользователями и экспертами в процессе создания, внедрения и использования систем искусственного интеллекта <b>умения:</b> применять методы распределения ролей в проектной команде, гибкие (agile) технологии выполнения проектных работ <b>навыки:</b> применения методов и средств коллективной работы, гибких (agile) технологии выполнения проектных работ в координации работ по созданию, внедрению и сопровождению систем искусственного интеллекта
3. ПК-5 Способен адаптировать и применять методы и алгоритмы машинного обучения для решения прикладных задач в различных предметных областях	ПК-5.1. Ставит задачи по адаптации или совершенствованию методов и алгоритмов для решения комплекса задач предметной области	<b>знания:</b> классы методов и алгоритмов машинного обучения <b>умения:</b> ставить задачи и адаптировать методы и алгоритмы машинного обучения <b>навыки:</b> использования методов и алгоритмов машинного обучения

## Раздел 2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОПОП

Дисциплина относится к части, формируемой участниками образовательных отношений ОПОП.

Дисциплина является обязательной

Для продолжения формирования заявленных компетенций необходимы знания предшествующих практик: Производственная практика. Организационно-управленческая (ПК-3), Производственная практика. Организационно-управленческая (ПК-4), Производственная практика. Организационно-управленческая (ПК-5)

Изучаемая дисциплина является основой для продолжения формирования указанных компетенций в следующих дисциплинах: Основы построения и эволюция систем искусственного интеллекта (ПК-3), Управление персоналом (продвинутый уровень) (ПК-4); практиках: Производственная практика (научно-исследовательская работа) (ПК-3), Производственная практика (научно-исследовательская работа) (ПК-4), Производственная практика (научно-исследовательская работа) (ПК-5); государственной итоговой аттестации в форме: Подготовка к процедуре защиты и защита выпускной квалификационной работы (ПК-3), Подготовка к процедуре защиты и защита выпускной квалификационной работы (ПК-4), Подготовка к процедуре защиты и защита выпускной квалификационной работы (ПК-4), Подготовка к процедуре защиты и защита выпускной квалификационной работы (ПК-4)

## Раздел 3. ОПИСАНИЕ ОБРАЗОВАТЕЛЬНЫХ ТЕХНОЛОГИЙ

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

подходы.

Основными стратегическими технологиями являются: исследовательские, лекционные занятия, практические занятия, процедуры самообучения

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

#### Раздел 4. СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

##### 3 семестр

Виды и темы занятий	Количество часов	Формируемые компетенции
<b>Раздел 1. Машинное обучение и интеллектуальный анализ данных</b>	<b>144</b>	ПК-3, ПК-4, ПК-5
Лекция. № 1. Машинное обучение как направление искусственного интеллекта 1. Понятие и виды машинного обучения. 2. Данные в машинном обучении. 3. Этапы машинного обучения.	2	
Практическое занятие. № 1. Инструментальные средства для машинного обучения	4	
Лекция. № 2. Подготовка данных для решения задач машинного обучения. 1. Предобработка данных: обработка пропусков, ошибок, явных и неявных дубликатов, фиктивных данных. 2. Поиск и обработка выбросов и аномалий в данных 3. Исследовательский анализ данных. 4. Статистический анализ данных.	2	
Практическое занятие. № 2. Решение задачи подготовки и предварительного анализа данных с использованием инструментальных средств 1. Предобработка данных. 2. Исследовательский и статистический анализ данных. 3. Визуализация данных.	4	
Лекция. № 3. Задача классификации в машинном обучении 1. Постановка задач классификации 2. Метрики, критерии качества в задачах классификации 3. Данные в задачах классификации 4. Методы решения задач классификации	2	
Практическое занятие. № 3. Алгоритмы машинного обучения для решения задач классификации 1. Метод k ближайших соседей 2. Логистическая регрессия 3. Решающие деревья	4	
Лекция. № 4. Задача регрессии в машинном обучении 1. Постановка задач регрессии 2. Метрики, критерии качества в задачах регрессии 3. Данные в задачах регрессии 4. Методы решения задач регрессии.	2	
Практическое занятие. № 4. Алгоритмы машинного обучения для решения задач регрессии 1. Метод k ближайших соседей 2. Линейная регрессия	4	

3. Решающие деревья	
Лекция. № 5. Задачи кластеризации и снижения размерности признаков пространства в машинном обучении 1. Обучение без учителя. Постановка задач кластеризации и снижения размерности признаков пространства 2. Метрики, критерии качества в задачах обучения без учителя 3. Методы решения задачи кластеризации 4. Методы решения задачи снижения признаков пространства	2
Практическое занятие. № 4. Алгоритмы машинного обучения в задачах обучения без учителя 1. Методы кластеризации: метод k средних, метод гауссовских смесей, DBSCAN 2. Методы снижения размерности признаков пространства: PCA, t-SNE	4
Лекция. № 6. Ансамблевые методы в машинном обучении 1. Задачи и принципы построения ансамблей 2. Виды ансамблей 3. Оценка неопределенности модели с помощью ансамблей	2
Практическое занятие. № 6. Алгоритмы ансамблевых методов в машинном обучении 1. Бэггинг 2. Стекинг 3. Бустинг	4
Задания для самостоятельной работы, в том числе выполнение Изучение лекционного материала, подготовка к текущему контролю, подготовка к лабораторным занятиям, изучение дополнительного материала, выполнение дополнительных заданий научно-исследовательского характера, самостоятельное решение задач для закрепления пройденного материала, подготовка к промежуточной аттестации	108
Иная контактная работа:	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.

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.	Станкевич, Лев Александрович. Интеллектуальные системы и технологии [Текст : Электронный ресурс] : учебник и практикум для вузов / Л. А. Станкевич. Москва: Юрайт, 2022. - 397 с ISBN 978-5-534-02126-4.	<a href="https://urait.ru/bcode/489694">https://urait.ru/bcode/489694</a>
2.	Остроух, А. В. Интеллектуальные информационные системы и технологии [Электронный ресурс] / Остроух А. В., Николаев А. Б. 3-е изд., стер. Санкт-Петербург: Лань, 2023. - 308 с. ISBN 978--5-507-48511-6.	<a href="https://e.lanbook.com/book/354536">https://e.lanbook.com/book/354536</a>
3.	Остроух, А. В. Системы искусственного интеллекта [Электронный ресурс] : монография / Остроух А. В., Суркова Н. Е.; Суркова Н. Е. 4-е изд., стер. Санкт-Петербург: Лань, 2024. - 228 с. ISBN 978-5-507-47478-3.	<a href="https://e.lanbook.com/book/379988">https://e.lanbook.com/book/379988</a>
4.	Мхитарян, Владимир Сергеевич. Статистика. В 2 ч. Часть 1 [Текст : Электронный ресурс] : учебник и практикум для вузов / В. С. Мхитарян [и др.] ; под редакцией В. С. Мхитаряна. 2-е изд. Москва: Юрайт, 2022. - 249 с ISBN 978-5-534-09353-7.	<a href="https://urait.ru/bcode/494854">https://urait.ru/bcode/494854</a>
5.	Мхитарян, Владимир Сергеевич. Статистика. В 2 ч. Часть 2 [Текст : Электронный ресурс] : учебник и практикум для вузов / В. С. Мхитарян, Т. Н. Агапова, С. Д. Ильенкова, А. Е. Суринов ; под редакцией В. С. Мхитаряна. 2-е изд. Москва: Юрайт, 2022. - 270 с ISBN 978-5-534-09357-5.	<a href="https://urait.ru/bcode/494855">https://urait.ru/bcode/494855</a>
6.	Ясницкий, Л. Н. Интеллектуальные системы [Электронный ресурс] : учебник / Ясницкий Л. Н. 2-е изд. Москва: Лаборатория знаний, 2020. - 224 с. ISBN 978-5-00101-897-1.	<a href="https://e.lanbook.com/book/151510">https://e.lanbook.com/book/151510</a>
<b>ЭЛЕКТРОННЫЕ ОБРАЗОВАТЕЛЬНЫЕ РЕСУРСЫ</b>		
1.	Научная электронная библиотека eLIBRARY.RU	<a href="http://elibrary.ru">http://elibrary.ru</a>
2.	Научная электронная библиотека «Киберленинка»	<a href="http://cyberleninka.ru">http://cyberleninka.ru</a>
<b>ПРОФЕССИОНАЛЬНЫЕ БАЗЫ ДАННЫХ И ИНФОРМАЦИОННЫЕ СПРАВОЧНЫЕ СИСТЕМЫ</b>		

1.	Справочно-правовая система Консультант+	<a href="http://www.consultant.ru">http://www.consultant.ru</a>
2.	Информационно-правовой портал Гарант	<a href="http://www.garant.ru">http://www.garant.ru</a>
3.	Профессиональные справочные системы Техэксперт	<a href="http://www.cntd.ru">http://www.cntd.ru</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.	306 (III)	Проектор мультимедийный Hitachi CP-X 5 (1), Экран настен. рулон. 200*200 Springroller Type D Medium (1), 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, STATISTICA 6.1 for Windows Russian
2.	335 (III)	Доска маркерная 120x240 см (1), Персональный компьютер Power RaY P550 (16), Экран настенный рулонный 200x200 см (1), 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, STATISTICA 6.1 for Windows Russian

## 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.

##### ПК-3:

1. Какой тип обучения включает в себя задачи, где алгоритм получает данные с метками и учится прогнозировать значения для новых данных?

- A) Обучение с учителем (Supervised Learning)
- B) Обучение без учителя (Unsupervised Learning)
- C) Полу-обучение (Semi-Supervised Learning)
- D) Обучение с подкреплением (Reinforcement Learning)

2. Какая метрика оценивает точность модели, учитывая и ложные положительные и ложные отрицательные результаты?

- A) Точность (Accuracy)
- B) Полнота (Recall)
- C) Точность и полнота (Precision and Recall)
- D) F1-мера (F1-Score)

3. Какой метод обучения подразумевает передачу знаний с уже обученной модели на новую модель для улучшения ее производительности?

- A) Обучение с подкреплением (Reinforcement Learning)
- B) Трансферное обучение (Transfer Learning)
- C) Обучение без учителя (Unsupervised Learning)
- D) Активное обучение (Active Learning)

#### **ПК-4:**

4. Какая техника позволяет преобразовать категориальные признаки в числовой формат для использования в моделях машинного обучения?

- A) One-Hot Encoding
- B) Перцептрон (Perceptron)
- C) Градиентный спуск (Gradient Descent)
- D) Ансамблевые методы

5. Какая величина НЕ является показателем центра распределения случайной величины?

- A) Математическое ожидание
- B) Медиана
- C) Дисперсия
- D) Мода

6. Какая величина является мерой разброса случайной величины относительно ее центра распределения:

- A) Математическое ожидание
- B) Размах вариации
- C) Дисперсия
- D) Эксцесс

#### **ПК-5:**

7. Какой алгоритм машинного обучения наиболее подходит для задачи обнаружения аномалий (аномалий в данных)?

- A) Линейная регрессия
- B) Метод опорных векторов (SVM)

C) К-средних (K-means)

D) Решающее дерево

8. Какой метод изучает влияние каждой переменной на выходной результат и используется в линейной регрессии?

A) PCA (Principal Component Analysis)

B) Градиентный спуск (Gradient Descent)

C) Анализ важности переменных

D) Сингулярное разложение (SVD)

9. Какой алгоритм используется для кластеризации данных, если заранее неизвестно количество кластеров?

A) К-средних (K-means)

B) DBSCAN (Density-Based Spatial Clustering of Applications with Noise)

C) Алгоритм Гаусса (Gaussian Mixture Model)

D) Иерархическая кластеризация (Hierarchical Clustering)

10. Какой метод уменьшает размерность данных, сохраняя при этом наибольшее количество информации?

A) Метод главных компонент (PCA)

B) t-Силуэт (t-SNE)

C) Линейная регрессия

D) К-средних (K-means)

11. Какой алгоритм обучения используется для прогнозирования вероятностей в задачах классификации?

A) Линейная регрессия

B) Дерево решений

C) Логистическая регрессия

D) Алгоритм k-ближайших соседей (k-NN)

12. Какой тип ансамблевого метода строит несколько деревьев решений и комбинирует их результаты для повышения точности?

A) Бэггинг (Bagging)

B) Бустинг (Boosting)

C) Случайный лес (Random Forest)

D) Градиентный бустинг (Gradient Boosting)

13. Какой алгоритм является основой для метода опорных векторов (SVM)?

A) Логистическая регрессия

B) Линейная регрессия

C) Метод главных компонент (PCA)

D) Классификатор опорных векторов

14. Какие методы регуляризации могут помочь в улучшении обобщающей способности модели?

A) L1 и L2 регуляризация

B) Стохастический градиентный спуск

C) Градиентный бустинг

D) Байесовская оптимизация

15. Какой из перечисленных методов оптимизации наиболее подходит для обучения моделей машинного обучения, когда данные сильно несбалансированы?

- A) Mini-batch градиентный спуск
- B) Adam оптимизатор
- C) Метод опорных векторов (SVM)
- D) Градиентный бустинг

Перечень вопросов для проведения промежуточной аттестации

### **ПК-3:**

#### **1. Введение в предмет машинного обучения.**

Основные задачи машинного обучения. Обучение с учителем и без учителя. Этапы построения модели машинного обучения. Обобщающая способность и переобучение. Метрики качества моделей машинного обучения. Функция потерь. Отложенная выборка. Кросс-валидация.

### **ПК-4:**

#### **2. Предобработка данных: работа с пропусками, дубликатами, ошибками в данных, поиск и анализ выбросов, кодирование и нормализация данных. Исследовательский анализ данных и конструирование признаков. Программные инструменты реализации методов машинного обучения.**

### **ПК-5:**

#### **3. Линейные модели и метрические методы.**

Особенности линейных моделей. Линейная регрессия и метод наименьших квадратов. Линейные модели в задачах классификации. Регуляризация L1, L2. Конструирование признаков и трансформирование данных. Логистическая регрессия. Метод ближайших  $k$  – соседей. Области применения линейных моделей: дисперсия и разброс.

#### **4. Решающие деревья.**

Особенности решающих деревьев. Решающие деревья в задачах классификации и регрессии. Обучение решающих деревьев. Критерий информативности. Критерий Джини и энтропийной информативности. Останов и «стрижка» деревьев.

#### **5. Случайный лес и градиентный бустинг.**

Основные недостатки решающих деревьев. Композиция алгоритмов. Рандомизация и композиция деревьев как защита от переобучения. Понятия случайного леса. Разложение ошибки: смещение, шум, разброс. Уменьшение корреляции базовых алгоритмов: бэггинг и метод случайных подпространств.

Основная идея бустинга. Бустинг на примере задачи регрессии. Описание алгоритма градиентного бустинга. Проблема переобучения градиентного бустинга и подбор гиперпараметров.

**6. Другие методы машинного обучения: Байесовские классификаторы и метод опорных векторов.**

Идея наивного Байесовского классификатора. Частотный и Байесовский вероятностный подход. Объединение вероятностных моделей. Задача спам-фильтра. Метод опорных векторов и его модификации. Особенности применения метода опорных векторов.

**7. Настройка моделей машинного обучения с размеченными данными.**

Гиперпараметры и проблема предмет ориентированных задач. Способы настройки гиперпараметров. Особенности оптимизации моделей машинного обучения. Основные методы оптимизации. Перебор по сетке. Случайный поиск. Эволюционные алгоритмы. Статистические методы.

**8. Обучение без учителя: кластеризация.**

Введение в кластеризацию. Основные методы кластеризации. Понижение размерности и отбор признаков. Матричные разложения. Типы кластерных структур. EM – Алгоритм и метод k – средних.

**9. Иерархическая кластеризация (таксономия).**

Агломеративная иерархическая кластеризация. Дендограммы и свойства монотонности. Свойства сжатия, растяжения и редуктивности. Примеры решения задач кластеризации с использованием агломеративной кластеризации.

