

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ  
ПОВОЛЖСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНОЛОГИЧЕСКИЙ УНИВЕРСИТЕТ



УТВЕРЖДАЮ  
Декан РТФ

УТВЕРЖДАЮ /А.Н. Дедов/  
(Ф.И.О. декана (директора института))

01.03.2023 г.

**РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)**

Б.1.1.19 Машинное обучение и анализ данных

*(код и наименование дисциплины по учебному плану)*

Направление подготовки  
(специальность)

27.03.04 Управление в технических системах

Квалификация выпускника

Бакалавр

*(бакалавр/магистр/специалист)*

Направленность

Управление и информатика в технических системах

Курс 2  
Семестр 4

**Распределение учебного времени**

Трудоемкость по учебному плану	144 / 4	часов/зачетных единиц
Лекции	36	часов
Лабораторные работы	36	часов
Практические занятия	-	часов
Иная контактная работа	-	часов
Всего контактной работы (без учета экз.)	72	часов
Контактная работа по экзамену	-	часов
Курсовой проект (работа)	-	семестр
Самостоятельная работа обучающихся (без учета экз.)	72	часов
Самостоятельная работа по подготовке к экзамену	-	часов
Экзамен	-	семестр
Зачет	-	семестр
БРК, ДЗ	4	семестр

                      
(год)

Программа составлена в соответствии с требованиями ФГОС ВО направления подготовки (специальности) 27.03.04 Управление в технических системах

Программу составили:

старший преподаватель	РТиС	СОГЛАСОВАНО	Н.А. Конкин
(должность)	(кафедра)		(И.О. Фамилия)

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

(наименование кафедры)		
16.01.2023	протокол №	8
(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.С. Буканова
		(И.О. Фамилия)

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

СОГЛАСОВАНО	А.Н. Дедов
	(И.О. Фамилия)

Эксперт(ы): Бастраков Александр Владиславович, заместитель главного инженера АО "ММЗ"

Рабочая программа проверена и зарегистрирована в УМЦ 06.03.2023 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. ОПК-6 Способен разрабатывать и использовать алгоритмы и программы, современные информационные технологии, методы и средства контроля, диагностики и управления, пригодные для практического применения в сфере своей профессиональной деятельности	ОПК-6.5 Реализует основные направления машинного обучения для решения задачи практического применения в сфере управления	<b>знания:</b> Знать математические методы решения задач классификации и регрессии, а также оптимизации для практического применения в сфере автоматизации управления <b>умения:</b> Уметь реализовывать алгоритмы математических методов и обработки соответствующих данных для решения задач практического применения в сфере автоматизации управления <b>навыки:</b> Навыки реализации программного кода, реализующего алгоритмы и математические методы для практического применения в сфере автоматизации управления
	ОПК-6.6 Проводит анализ эффективности алгоритмов машинного обучения, делает выбор и реализует модель обучения для решения задачи практического применения в сфере управления	<b>знания:</b> Знать способы оценки эффективности и настройки алгоритмов для практического применения в сфере автоматизации управления <b>умения:</b> Уметь реализовывать оценку эффективности и настройки алгоритмов для практического применения в сфере автоматизации управления <b>навыки:</b> Навыки реализации программного кода для оценки эффективности и настройки алгоритмов в сфере автоматизации управления
	ОПК-6.7 Использует соответствующий математический аппарат для разработки моделей, пригодных для практического применения в сфере своей профессиональной деятельности	<b>знания:</b> Знать основные численные и аналитические математические методы для практического применения в инженерно-профессиональной сфере <b>умения:</b> Умения в части использования математического аппарата для разработки эффективных моделей в инженерно-профессиональной сфере <b>навыки:</b> Знания в выборе и применении программных средств для разработки эффективных моделей в инженерно-профессиональной сфере

## Раздел 2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОПОП

Дисциплина относится к обязательной части ОПОП.

Дисциплина является обязательной

Для продолжения формирования заявленных компетенций необходимы знания предшествующих дисциплин: Алгоритмизация и программирование (ОПК-6)

Изучаемая дисциплина является основой для продолжения формирования указанных

компетенций в следующих государственной итоговой аттестации в форме: Выполнение, подготовка к процедуре защиты и защита выпускной квалификационной работы (ОПК-6)

### Раздел 3. ОПИСАНИЕ ОБРАЗОВАТЕЛЬНЫХ ТЕХНОЛОГИЙ

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

Основными стратегическими технологиями являются: лекционные занятия, практические и лабораторные занятия

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

### Раздел 4. СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

#### 4 семестр

Виды и темы занятий	Количество часов	Формируемые компетенции
<b>Введение в машинное обучение и анализ данных</b>	<b>48</b>	ОПК-6
Лекция. Основы функционального анализа: метрические и нормированные пространства. Линейная алгебра в машинном обучении	2	
Лекция. Теория вероятностей и математическая статистика базовые инструменты анализа данных.	4	
Лекция. Основы анализа и численные методы оптимизации в задачах машинного обучения.	2	
Лекция. Введение в предмет машинного обучения.	4	
Лабораторная работа. Изучение основ подготовки данных в задачах машинного обучения	12	
Задания для самостоятельной работы, в том числе выполнение Проработка теоретического материала по темам курса, выполнение контрольных работ	24	
<b>Подходы к решению задач классификации, кластеризации</b>	<b>48</b>	ОПК-6
Лекция. Линейные модели и метрические методы.	4	
Лекция. Решающие деревья. Случайный лес и градиентный бустинг.	4	
Лекция. Другие методы машинного обучения: Байесовские классификаторы и метод опорных векторов.	4	
Лабораторная работа. Изучение методов классификации и кластеризации данных	12	
Задания для самостоятельной работы, в том числе выполнение Проработка теоретического материала по темам курса, выполнение контрольных работ	24	
<b>Глубокое обучение</b>	<b>48</b>	ОПК-6
Лекция. Нейронные сети – универсальные классификаторы. Сверточные ИНС – глубокое обучение.	4	
Лекция. Настройка моделей машинного обучения с размеченными данными.	4	
Лекция. Обучение без учителя: кластеризация. Иерархическая кластеризация (таксономия).	4	
Лабораторная работа. Изучение структур ИНС и синтез	12	

моделей на основе ИНС		
Задания для самостоятельной работы, в том числе выполнение Проработка теоретического материала по темам курса, выполнение контрольных работ	24	
Иная контактная работа: дифференцированный зачет (БРК)	0	

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

Изучение дисциплины машинное обучение и анализ данных 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 машинное обучение и анализ, 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 машинное обучение и анализ. Содержание **самостоятельной работы** определяется рабочей программой дисциплины машинное обучение и анализ, 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 машинное обучение и анализ, к электронным библиотечным системам, профессиональным базам данных и информационным справочным системам. Изучение дисциплины машинное обучение и анализ включает выполнение лабораторных работ. Периодичность проведения, формы текущего контроля успеваемости, система оценивания хода освоения дисциплин представлены в рабочей программе. Формой промежуточной аттестации по дисциплине машинное обучение и анализ данных бально-рейтинговый контроль.

## 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.	Векторная алгебра и аналитическая геометрия [Текст] : сб. контрольных работ для студентов специальностей 291400, 550100, 291500, 290300 / [сост. Н. К. Томилова]. Йошкар-Ола: МарГТУ, 2007. - 34 с. Экземпляры: всего 5.	5
2.	Годин, Александр Михайлович. Статистика [Текст] :	60

	[учеб. для вузов по экон. специальностям и направлениям] / А. М. Годин. 3-е изд., перераб. М.: Дашков и К, 2005. - 470 с. ISBN 5-94798-504-7. Экземпляры: всего 60.	
3.	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
4.	Круглов, Владимир Васильевич. Искусственные нейронные сети [Текст] : теория и практика / В. В. Круглов, В. В. Борисов. М.: Горячая линия - Телеком, 2001. - 381 с. ISBN 5-93517-031-0. Экземпляры: всего 5.	5
5.	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>
6.	Афонин, В. Л. Интеллектуальные робототехнические системы [Электронный ресурс] / Афонин В. Л., Макушкин В. А. 2-е изд. Москва: ИНТУИТ, 2016. - 222 с. ISBN 5-9556-00024-8.	<a href="https://e.lanbook.com/book/100607">https://e.lanbook.com/book/100607</a>
7.	Войтович, И. Д. Интеллектуальные сенсоры [Электронный ресурс] / Войтович И. Д., Корсунский В. М. 2-е изд. Москва: ИНТУИТ, 2016. - 1164 с. ISBN 978-5-9963-0124-9.	<a href="https://e.lanbook.com/book/100608">https://e.lanbook.com/book/100608</a>
8.	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>
9.	Селянкин, В. В. Компьютерное зрение. Анализ и обработка изображений [Электронный ресурс] : учебное пособие для вузов / Селянкин В. В. 3-е изд., стер. Санкт-Петербург: Лань, 2023. - 152 с. ISBN 978-5-507-45583-6.	<a href="https://e.lanbook.com/book/276455">https://e.lanbook.com/book/276455</a>
10.	Шапиро, Л. Компьютерное зрение [Электронный ресурс] / Шапиро Л., Стокман Д. 4-е изд. Москва: Лаборатория знаний, 2020. - 763 с. ISBN 978-5-00101-696-0.	<a href="https://e.lanbook.com/book/135496">https://e.lanbook.com/book/135496</a>
ЭЛЕКТРОННЫЕ ОБРАЗОВАТЕЛЬНЫЕ РЕСУРСЫ		
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>

## 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.	414 (III)	Монитор 19" Samsung 940N (KSB) TFT Silver. Round Simple (6), Персональный компьютер 6 Atlant A2X2/2G(3)/монитор Viewsonic VA2013wm/3Y (5), ПК S404,2	Microsoft Windows Enterprise, Microsoft Office Standard, Агент Dr.Web, Microsoft Access, Microsoft Visio

		400W/Intel Core i3 540/клав.,мышь,монит. 21,5" VA2248-LED (3), Систем.блок Athlon 64 3500/512Mb*2/160Gb/FDD/DVD-RW клав.мышь.ковр. (6), Комплект учебной мебели (1)	Professional, Microsoft Project Professional, Microsoft Visual Studio Enterprise, Комплект ПО для решения основных пользовательских задач
2.	416 (III)	Монитор 19"Samsung 943N(KSB) TFT (1), Отладочный стенд для макетирования и проверки эл.механических систем (1), Отладочный стенд для синтеза эл.механических систем с микроконтроллерным управл. (1), Проектор мультимедийный Hitachi CP-RX93 (1), Систем.блок Core 2DUO E6320/1024Mb*2/160Gb/GF8500GT/DVD-RW/FDD клав.мышь.коврик (1), Учебно-лабораторный стенд "Датчики технологических параметров" ДТП1-Н-Р (1), Учебно-лабораторный стенд "Исполнительный двигатель постоянного тока" ИДПТ1-Н-Р (1), Учебно-лабораторный стенд "Сервопривод" СПМ1-С-К (1), Учебно-лабораторный стенд датчиков скорости вращения (1), Учебно-лабораторный стенд исполнительного шагового двигателя ИШД1-Н-К (1), Учебно-лабораторный стенд по разделам дисциплин "Автоматика на основе программируемого контроллера" (1), Учебно-лабораторный стенд по разделу дисциплин "Автоматизация технологических процессов и производств на основе приборов" АТПП2-Н-К (1), Учебно-лабораторный стенд по разделу дисциплин "Автоматизация технологических процессов и производств" АТПП1-Н-К (1), Учебно-лабораторный стенд по разделу дисциплин "Основы автоматизации производства" ОАП1-Н-Р (1), Учебно-лабораторный стенд по разделу дисциплин "Средства автоматизации и управления" САУ1-Н-К (1),	Microsoft Windows Enterprise, Microsoft Office Standard, Агент Dr.Web, 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.

Искусственные нейронные сети (ИНС) — модели машинного обучения, использующие комбинации распределенных простых операций, зависящих от обучаемых параметров, для обработки входных данных. Какого вида ИНС не существует?

1. Рекуррентные
2. Противоборствующие
3. Импульсные
4. Наивные

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

1. Классификация
2. Кластеризация
3. Переобучение
4. Регрессия

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?

1. Цвет
2. Форма
3. Текстуры
4. Глубина, количество пикселей

Особых успехов нейросети достигли в работе с изображениями. Но что из этого нейросети не могут сделать?

1. Стилизовать вашу фотографию под работу импрессиониста
2. Догадаться, что вы нарисовали
3. Омолаживать и состаривать лица на фотографиях
4. Пластическую коррекцию лица

Кто создал первую модель искусственных нейронных сетей?

- 1.Мак-Каллок и Питтс
- 2.Дэвид И. Румельхарт, Дж. Е. Хинтон и Рональд Дж. Вильямс
- 3.Фрэнк Розенблатт
- 4.Ян Лекун

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

- 1.Обучение без учителя
- 2.Обучение с подкреплением
- 3.Обучение с учителем
- 4.Глубинное обучение

Допустим, нам нужно рассчитать необходимые параметры для создания обшивки самолета. Какая из областей машинного обучения нам в этом пригодится?

- 1.Латентная модель
- 2.Компьютерное зрение
- 3.Обучение ранжированию
- 4.Предсказательное моделирование

Что такое машинное обучение?

- А) Программирование компьютеров
- В) Процесс, в ходе которого компьютеры обучаются из опыта

- C) Ремонт роботов
- D) Работа с металлом

Какие из перечисленных видов обучения существуют в машинном обучении?

- A) Обучение с учителем
- B) Обучение без учителя
- C) Обучение с подкреплением
- D) Все вышеперечисленные

Что такое алгоритм классификации в машинном обучении?

- A) Процесс присвоения категории объекту на основе его характеристик
- B) Пай-рецепт
- C) Математическая формула для расчета расстояния между точками
- D) Название нового видеоигрового жанра

Какие виды данных используются в анализе данных в робототехнике?

- A) Только числовые данные
- B) Только текстовые данные
- C) Числовые, текстовые, изображения и другие типы данных
- D) Только графические данные

Что такое нейронные сети?

- A) Сети нейронов в мозге
- B) Математические модели, имитирующие работу нейронов в мозге
- C) Только название науки
- D) Программы для создания графики

Какие алгоритмы используются в обучении с учителем?

- A) Кластеризация
- B) Линейная регрессия
- C) Метод опорных векторов
- D) A и B

Что такое обучение без учителя?

- A) Обучение, при котором компьютеры учатся сами без учителя
- B) Обучение без использования учителя или меток в данных
- C) Обучение с использованием роботов
- D) Обучение с использованием мультимедийных материалов

## Перечень вопросов для проведения промежуточной аттестации

### 1. Основы функционального анализа: метрические и нормированные пространства.

Понятие метрики. Аксиомы метрики. Типы метрик (Евклидова, Хемминга, Равномерного приближения, косинусное расстояние). Нормированные пространства. Аксиомы нормы.

### 2. Линейная алгебра в машинном обучении.

Объекты машинного обучения: скаляр, вектор, матрица и тензор. Матричные операции. Ранг матрицы и определитель. Линейная независимость. Теорема Кронекера-Капелли. Типы матриц и понятие собственного вектора. Сингулярное разложение матрицы.

### 3. Теория вероятностей и математическая статистика базовые инструменты анализа данных.

Понятие случайности. Свойства случайной величины. Условная вероятность и теорема Байеса. Парадокс

Монти-Холла. Непрерывность и дискретность случайной величины. Основные законы распределения случайных величин. Гомоскедастичность и гетероскедастичность. Основные статистики в машинном обучении. Центральная предельная теорема. Метод максимального правдоподобия. Доверительные интервалы.

#### **4. Основы анализа и численные методы оптимизации в задачах машинного обучения.**

Функции многих переменных и понятие градиента. Теорема Вейерштрасса (необходимые и достаточные условия экстремума). Оптимизация градиентным методом. Матрицы Якоби и Гессе. Оптимизация с ограничениями. Эвристические методы оптимизации (генетический алгоритм, метод отжига и Нелдера-Мида). Линейный метод наименьших квадратов.

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

Основные задачи машинного обучения. Особенности построения модели машинного обучения. Обобщающая способность и переобучение. Теорема о бесплатном завтраке. Обучение с учителем и без учителя. Отложенная выборка. Кросс-валидация. Подготовка данных: масштабирование, биннинг, кодирование. Программные инструменты реализации методов машинного обучения.

#### **6. 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.

#### **7. 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.

#### **8. 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.

#### **9. 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.

#### **10. Нейронные сети – универсальные классификаторы.**

Перцептрон Розенблатта – прообраз современных моделей биологически активного нейрона. Проблема исключающего или - XOR. Геометрический смысл скрытого слоя. Функции активации и другие особенности построения топологии искусственной нейронной сети (ИНС). Алгоритм обратного распространения ошибки. Современные типы ИНС. Борьба с переобучением.

#### **11. Сверточные ИНС – глубокое обучение.**

Сверточные ИНС – дальнейшее развитие ИНС. Понятие свертки и особенности работы сверточной ИНС. Операция субдискретизации и работа свертки. Сверточные ИНС для задач машинного зрения и распознавания звуковых команд. Перенос обучение и использование предобученной ИНС.

#### **Лекция 12. 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.

### **13. 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.

### **14. Иерархическая кластеризация (таксономия).**

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