

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ  
РОССИЙСКОЙ ФЕДЕРАЦИИ  
Федеральное государственное бюджетное образовательное учреждение  
высшего образования  
«ДАГЕСТАНСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»  
Факультет информатики и информационных технологий

## **РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ**

### **Методы машинного обучения**

Кафедра информационных систем и технологий программирования  
факультета информатики и информационных технологий

**Образовательная программа магистратуры**  
09.04.03 Прикладная информатика

**Направленность (профиль) программы**  
Разработка и внедрение информационных систем  
Цифровая экономика

**Форма обучения**  
очная, заочная

**Статус дисциплины:**  
входит в обязательную часть ОПОП

Махачкала, 2022

Рабочая программа дисциплины «Методы машинного обучения» составлена в 2022 году в соответствии с требованиями ФГОС ВО – магистратура по направлению подготовки 09.04.03 Прикладная информатика от «19» сентября 2017г. № 916.

Разработчик(и): кафедра информационных систем и технологий программирования, Магомедгаджиев Ш.М., к.э.н., доцент

Рабочая программа дисциплины одобрена:  
на заседании кафедры ИСиТП от «1» марта 2022г., протокол № 8

Зав. кафедрой  Исмиханов З.Н.  
(подпись)

на заседании Методической комиссии факультета ИиИТ  
от «17» марта 2022г., протокол № 7

Председатель  Бакмаев А.Ш.  
(подпись)

Рабочая программа дисциплины согласована с учебно-методическим управлением «31» марта 2022г.

Начальник УМУ  Гасангаджиева А.Г.  
(подпись)

## Аннотация рабочей программы дисциплины

Дисциплина «Методы машинного обучения» входит в обязательную часть образовательной программы магистратуры по направлению 09.04.03 Прикладная информатика.

Дисциплина реализуется на факультете информатики и информационных технологий кафедрой информационных систем и технологий программирования.

Содержание дисциплины охватывает круг вопросов, связанных с изучением видов машинного обучения – кластеризации, классификации, регрессии и особенностей его применения для решения прикладных задач.

Дисциплина нацелена на формирование следующих компетенций выпускника: универсальных - УК-1; общепрофессиональных ОПК -3, ОПК-7.

Преподавание дисциплины предусматривает проведение следующих видов учебных занятий: лекции, практические занятия, лабораторные занятия, самостоятельная работа.

Рабочая программа дисциплины предусматривает проведение следующих видов контроля успеваемости: текущий контроль в форме опроса, тестов, контрольных работ и промежуточный контроль в форме экзамена.

Объем дисциплины 5 зачетных единиц, в том числе 180 в академических часах по видам учебных занятий

форма обучения - очная

Семестр	Учебные занятия					СРС, в том числе экзамен	Форма промежуточной аттестации
	в том числе:						
	всего	Контактная работа обучающихся с преподавателем			экзамен		
		всего	Лекции	Лабораторные занятия			
5	180	36	18	18		144	экзамен

форма обучения - заочная

Семестр	Учебные занятия					СРС, в том числе экзамен	Форма промежуточной аттестации
	в том числе:						
	всего	Контактная работа обучающихся с преподавателем			экзамен		
		всего	Лекции	Лабораторные занятия			
5	180	10	4	6		170	экзамен

## 1. Цели освоения дисциплины

Целями освоения дисциплины «Методы машинного обучения» является приобретение обучающимися теоретических знаний и практических навыков применения основных методов машинного обучения для задач классификации, кластеризации и регрессии (прогнозирования).

Задачи курса:

- сформировать теоретические знания по основам машинного обучения;
- научиться применять методы машинного обучения для решения прикладных задач в различных областях профессиональной деятельности;
- научиться применять навыки программирования на высокоуровневом языке Python для решения задач машинного обучения.

## 2. Место дисциплины в структуре ОПОП магистратуры

Дисциплина «Методы машинного обучения» входит в обязательную часть образовательной программы по направлению 09.04.03 Прикладная информатика.

При изучении дисциплины «Методы машинного обучения» предполагается, что студент владеет основами матричной алгебры, математического анализа, теории вероятностей, программирования, экономической теории, экономики и статистики фирмы в объёме, предусмотренным ФГОС ВО подготовки магистров.

Данную учебную дисциплину дополняет параллельное или последующее освоение следующих дисциплин: «Современные технологии разработки программного обеспечения», «Нейронные сети» и «Интеллектуальные ИС и методы искусственного интеллекта», «Технологии создания информационно-аналитических систем»

## 3. Компетенции обучающегося, формируемые в результате освоения дисциплины (перечень планируемых результатов обучения и процедура освоения).

Код и наименование компетенции из ОПОП	Код и наименование индикатора достижения компетенций (в соответствии с ОПОП)	Планируемые результаты обучения	Процедура освоения
УК-1. Способен осуществлять критический анализ проблемных ситуаций на основе системного подхода, вырабатывать стратегию действий	М-ИУК-1.1. Анализирует проблемную ситуацию как систему, выявляя ее составляющие и связи между ними М-ИУК-1.2. Определяет пробелы в информации, необходимой для решения проблемной ситуации, и проектирует процессы по их устранению М-ИУК-1.3. Критически оценивает надежность источников информации, работает с противоречивой информацией из разных источников М-ИУК-1.4. Разрабатывает и содержательно аргументирует стратегию решения проблемной ситуации на основе системного и междисциплинарного подходов	<i>Знает:</i> процедуры критического анализа, методика анализа результатов исследования и разработки стратегий проведения исследований, организации процесса принятия решения. <i>Умеет:</i> принимать конкретные решения для повышения эффективности процедур анализа проблем, принятия решений и разработки стратегий. <i>Владеет:</i> методами установления причинно - следственных связей и определения наиболее значимых среди них; методиками постановки цели и определения способов ее достижения; методиками разработки стратегий, действий при проблемных	Опрос, тестирование, контрольная работа

	М-ИУК-1.5 Строит сценарии реализации стратегии, определяя возможные риски и предлагая пути их устранения	ситуациях	
ОПК-3. Способен анализировать профессиональную информацию, выделять в ней главное, структурировать, оформлять и представлять в виде аналитических обзоров с обоснованными выводами и рекомендациями	ОПК-3.1. Знать принципы, методы и средства анализа и структурирования профессиональной информации; ОПК-3.2. Уметь анализировать профессиональную информацию, выделять в ней главное, структурировать, оформлять и представлять в виде аналитических обзоров	<i>Знает:</i> принципы, методы и средства анализа и структурирования профессиональной информации; <i>Умеет:</i> анализировать профессиональную информацию, выделять в ней главное, структурировать, оформлять и представлять в виде аналитических обзоров;	Опрос, тестирование, контрольная работа
ОПК-7. Способен использовать методы научных исследований и математического	ОПК-7.1. Знать логические методы и приемы научного исследования; методологические принципы современной науки, направления, концепции, моделирования в области проектирования и управления информационными системами источники знания и приемы работы с ними; основные особенности научного метода познания; программно-целевые методы решения научных проблем; основы моделирования управленческих решений; динамические оптимизационные модели; математические модели оптимального управления для непрерывных и дискретных процессов, их сравнительный анализ; многокритериальные методы принятия решений; ОПК-7.2. Уметь осуществлять методологическое обоснование научного исследования;	<i>Знает:</i> логические методы и приемы научного исследования; методологические принципы современной науки, направления, концепции, моделирования в области проектирования и управления информационными системами источники знания и приемы работы с ними; основные особенности научного метода познания; программно-целевые методы решения научных проблем; основы моделирования управленческих решений; динамические оптимизационные модели; математические модели оптимального управления для непрерывных и дискретных процессов, их сравнительный анализ; многокритериальные методы принятия решений; <i>Умеет:</i> осуществлять методологическое обоснование научного исследования;	Опрос, тестирование, контрольная работа

#### 4. Объем, структура и содержание дисциплины.

4.1. Объем дисциплины составляет 5 зачетных единиц, 180 академических часов.

4.2. Структура дисциплины.

4.2.1. Структура дисциплины

Очная форма

№ п/п	Разделы и темы дисциплины	Семестр	Неделя семестра	Виды учебной работы, включая самостоятельную работу студентов и трудоемкость (в часах)				Самостоятельная работа	Формы текущего контроля успеваемости (по неделям семестра) Форма промежуточной аттестации (по семестрам)
				Лекции	Практические занятия	Лабораторные занятия	Контроль самост. раб.		
Модуль 1. Основы машинного обучения									
1	Введение в методы машинного обучения	1	1-4	4		4		28	Опрос, тестирование, контрольная работа
	<i>Итого по модулю 1:</i>			4		4		28	
Модуль 2. Введение в регрессионный анализ									
1	Метод регрессии и метод главных компонент	1	5-7	6		2		28	Опрос, тестирование, контрольная работа
	<i>Итого по модулю 2:</i>			6		2		28	Зачет
Модуль 3. Методы кластеризации									
1	Методы кластеризации в машинном обучении	1	8-10	4		6		26	Опрос, тестирование, контрольная работа
2	<i>Итого по модулю 3:</i>	1	11-13	4		6		26	
Модуль 4. Методы классификации									
	Методы классификации в машинном обучении		14-18	4		6		26	Опрос, тестирование, контрольная работа
	<i>Итого по модулю 4:</i>	1		4		6		26	
Модуль 4. Подготовка к экзамену									
	<i>Итого по модулю 4:</i>							36	экзамен
	<b>ИТОГО:</b>			18		18		144	

заочная форма

№ п/п	Разделы и темы дисциплины	Семестр	Неделя семестра	Виды учебной работы, включая самостоятельную работу студентов и трудоемкость (в часах)				Самостоятельная работа	Формы текущего контроля успеваемости (по неделям семестра) Форма промежуточной аттестации (по семестрам)
				Лекции	Практические занятия	Лабораторные занятия	Контроль самост. раб.		

Модуль 1. Основы машинного обучения								
1	Введение в методы машинного обучения	1	1-4	1	2		33	Опрос, тестирование, контрольная работа
	<i>Итого по модулю 1:</i>			1	2		33	
Модуль 2. Введение в регрессионный анализ								
1	Метод регрессии и метод главных компонент	1	5-7	1	1		34	Опрос, тестирование, контрольная работа
	<i>Итого по модулю 2:</i>			1	1		34	Зачет
Модуль 3. Методы кластеризации								
1	Методы кластеризации в машинном обучении	1	8-10	1	2		33	Опрос, тестирование, контрольная работа
2	<i>Итого по модулю 3:</i>	1	11-13	1	2		33	
Модуль 4. Методы классификации								
	Методы классификации в машинном обучении		14-18	1	1		34	
	<i>Итого по модулю 4:</i>	1		1	1		34	
Модуль 4. Подготовка к экзамену								
	<i>Итого по модулю 4:</i>						36	экзамен
	<b>ИТОГО:</b>			4	6		170	

#### 4.3. Содержание дисциплины, структурированное по темам (разделам).

##### 4.3.1. Содержание лекционных занятий по дисциплине.

###### Лекционный курс

№ п/п	Наименование темы	Трудоемкость	Содержание	Формируемые компетенции	Результаты освоения (знать, уметь, владеть)	Технологии обучения
1.	Введение в методы машинного обучения	4	Виды обучения: с учителем, без учителя, с подкреплением. Основные типы задач: задача классификации, задача регрессии, задача кластеризации, задача прогнозирования, задача ранжирования. Библиотеки языка Python - NumPy, Scikit-Learn.	УК-1. ОПК-3. ОПК-7.	Знать математические основы теории машинного обучения; основные классы алгоритмов машинного обучения и их представителей и их взаимосвязь, достоинства и недостатки; Знать инструментарий Python и библиотеки Sklearn для предобработки данных и разработки алгоритмов машинного обучения; Знать возможности актуальных алгоритмов машинного обучения, которые широко используются на практике в научных исследованиях, основные сферы их применения	Опрос, тестирование, контрольная работа
2.	Метод регрессии и метод главных компонент	2	Метод наименьших квадратов. Модификации метода наименьших	УК-1. ОПК-3.	Знать теоретические основы машинного обучения, метод регрессии и метод главных компонент.	Опрос, тестирование, контрольная работа

			квадратов. Понижение размерностей. Сингулярное разложение. Метод главных компонент.		Уметь решать типовые задачи регрессии и метода главных компонент и формулировать прикладные задачи в терминах машинного обучения.	я работа, кейс-задача
3.	Методы кластеризации в машинном обучении	2	Метрические кластеризация. метод k-средних (k-means), их модификации. Иерархическая кластеризация. Деревья решений, алгоритм случайного леса.	УК-1. ОПК-3.	Знать теоретические основы машинного обучения, методы кластеризации. Уметь решать типовые задачи кластеризации и формулировать прикладные задачи в терминах машинного обучения.	Опрос, тестирование, контрольная работа
4.	Методы классификации в машинном обучении	2	Методы классификации. Методы наибольшего правдоподобия. Метод kNN. Наивный Байесовский классификатор. Решающие деревья и случайный лес. Логистическая регрессия. Метод опорных векторов.	УК-1. ОПК-3.	Знать теоретические основы машинного обучения, методы классификации. Уметь решать типовые задачи классификации и формулировать прикладные задачи в терминах машинного обучения.	Опрос, тестирование, контрольная работа

### Лабораторные занятия

№ п/п	Наименование темы	Трудоемкость	Содержание	Формируемые компетенции	Результаты освоения (знать, уметь, владеть)	Технологии обучения
1.	Введение в методы машинного обучения	2	Лабораторная работа №1: Основы анализа данных на Python	ОПК-7	Владеть основами анализа данных на Python для разработки алгоритмов машинного обучения	Опрос, кейс-задача
2.	Введение в методы машинного обучения	2	Лабораторная работа №2: Библиотека Sklearn и предобработка данных	ОПК-7	Владеть инструментарием библиотеки Sklearn для предобработки данных и разработки алгоритмов машинного обучения	Опрос, кейс-задача
3.	Метод регрессии и метод главных компонент	2	Лабораторная работа №3: Оценка линейной регрессии	ОПК-7	Владеть инструментарием Python, библиотеки Sklearn для разработки алгоритмов линейной регрессии (Linear Regression)	Опрос, кейс-задача
4.	Методы кластеризации в машинном обучении	2	Лабораторная работа №4: Композиции моделей градиентный бустинг и блендинг	ОПК-7	Владеть инструментарием Python, библиотеки Sklearn для разработки алгоритмов кластеризации (Gradient Boosting Classifier)	Опрос, кейс-задача
5.	Методы кластеризации в машинном обучении	2	Лабораторная работа №6: Кластеризация с помощью k-средних	ОПК-7	Владеть инструментарием Python, библиотеки Sklearn для разработки алгоритмов кластеризации с помощью k-средних (KMeans)	
6.	Методы классификации	2	Лабораторная работа №5:	ОПК-7	Владеть инструментарием Python, библиотеки Sklearn	Опрос, кейс-задача



	в машинном обучении.		Деревья решений		для разработки алгоритмов дерева решений (Decision Tree Classifier)	
7.	Методы классификации в машинном обучении	2	Лабораторная работа №7: Метод k-ближайших соседей	ОПК-7	Владеть инструментарием Python, библиотеки Sklearn для разработки алгоритмов классификации с помощью k- ближайших соседей (K-Nearest Neighbors)	Опрос, кейс-задача
8.	Методы классификации в машинном обучении	2	Лабораторная работа №8: Логистическая регрессия	ОПК-7	Владеть инструментарием Python, библиотеки Sklearn для разработки алгоритмов логистической регрессии (Logistic Regression)	Опрос, кейс-задача
9.	Методы классификации в машинном обучении	2	Лабораторная работа №9: Бинарная линейная классификация	ОПК-7	Владеть инструментарием Python, библиотеки Sklearn для разработки алгоритмов бинарной линейной классификации (Logistic Regression)	Опрос, кейс-задача

## Модуль 1. Основы машинного обучения

### Тема 1. Введение в методы машинного обучения

Виды обучения: с учителем, без учителя, с подкреплением. Основные типы задач: задача классификации, задача регрессии, задача кластеризации, задача прогнозирования, задача ранжирования. Основные проблемы машинного обучения: недостаточный объем обучающей выборки, пропуски в данных, переобучение. Обзор основных необходимых библиотек языка Python. Библиотека NumPy. Знакомство с библиотекой машинного обучения Scikit-Learn.

## Модуль 2. Введение в регрессионный анализ

### Тема 2. Метод регрессии и метод главных компонент.

Метод наименьших квадратов. Модификации метода наименьших квадратов. Одномерная линейная регрессионная модель. Одномерная нелинейная регрессионная модель. Многомерная линейная регрессионная модель. Понижение размерностей. Сингулярное разложение. Метод главных компонент. Многомерная нелинейная регрессионная модель.

## Модуль 3. Методы кластеризации

### Тема 3. Методы кластеризации в машинном обучении

Метрические кластеризация. метод k-средних (k-means), их модификации, метод k-мод (k-modes), методы нечеткой кластеризации с примерами (c-means). Иерархическая кластеризация. Метод формального элемента, логические дендрограммы. Коэффициент корреляции Пирсона и ранговый коэффициент корреляции Кендалла. Корреляционные плеяды. Деревья решений, алгоритм случайного леса.

## Модуль 4. Методы классификации

### Тема 4. Методы классификации в машинном обучении

Методы классификации. Методы наибольшего правдоподобия. Метод kNN. Наивный Байесовский классификатор. Решающие деревья и случайный лес. Нормированные векторные пространства. Нормы и нормированные векторные пространства. Метрические и нормированные пространства. Классификация в векторных пространствах. Логистическая регрессия. Метод опорных векторов.

### 4.3.2. Содержание лабораторно-практических занятий по дисциплине.

#### Лабораторные работы (лабораторный практикум)

1. Лабораторная работа №1: Основы анализа данных на Python
2. Лабораторная работа №2: Библиотека Sklearn и предобработка данных
3. Лабораторная работа №3: Оценка линейной регрессии
4. Лабораторная работа №4: Композиции моделей градиентный бустинг и блендинг
5. Лабораторная работа №6: Кластеризация с помощью k-средних
6. Лабораторная работа №5: Деревья решений
7. Лабораторная работа №7: Метод k-ближайших соседей
8. Лабораторная работа №8: Логистическая регрессия
9. Лабораторная работа №9: Бинарная линейная классификация

#### 5. Образовательные технологии

Традиционные образовательные технологии - лекции в сочетании с практическими занятиями, семинарами и с лабораторными работами, самостоятельное изучение определенных разделов. Использование персональных компьютеров при выполнении лабораторных работ и сдаче итогового экзамена. Чтение лекций с использованием компьютера и проектора, проведение лабораторных работ в компьютерном классе.

При реализации учебной дисциплины используются электронные практикумы, электронные учебники, презентации средства диагностики и контроля, разработанные специалистами кафедры т.д.

Удельный вес занятий, проводимых в интерактивной форме, составляет 20% аудиторных занятий.

#### 6. Учебно-методическое обеспечение самостоятельной работы студентов.

*Форма контроля и критерий оценок*

В соответствии с учебным планом предусмотрен экзамен в четвертом семестре.

Формы контроля: текущий контроль, промежуточный контроль по модулю, итоговый контроль по дисциплине предполагают следующее распределение баллов.

Текущий контроль

- Выполнение 1 домашней работы 10 баллов
- Активность в системе Moodle 10 баллов

Промежуточный контроль

Примерное распределение времени самостоятельной работы студентов

Вид самостоятельной работы	Примерная трудоёмкость, а.ч.		Формируемые компетенции
	очная	заочная	
<b>Текущая СРС</b>			
работа с лекционным материалом, с учебной литературой	16	18	УК-1
опережающая самостоятельная работа (изучение нового материала до его изложения на занятиях)	8	10	УК-1
самостоятельное изучение разделов дисциплины	32	36	УК-1
выполнение домашних заданий, домашних контрольных работ	24	28	УК-1
подготовка к лабораторным работам, к практическим и семинарским занятиям	12	12	УК-1ОПК-3, ОПК-7
подготовка к контрольным работам, коллоквиумам, зачётам	16	18	УК-1, ОПК-3, ОПК-7
<b>Творческая проблемно-ориентированная СРС</b>			
выполнение расчётно-графических работ	16	18	УК-1

поиск, изучение и презентация информации по заданной проблеме, анализ научных публикаций по заданной теме	8	10	УК-1
исследовательская работа, участие в конференциях, семинарах, олимпиадах	10	10	ОПК-7
анализ данных по заданной теме, выполнение расчётов, составление схем и моделей на основе собранных данных	10	10	ОПК-7
<b>Итого СРС:</b>	<b>152</b>	<b>170</b>	

Самостоятельная работа студентов (СРС) включает контролируемую и внеаудиторную самостоятельную работу, направлена на повышение качества обучения, углубление и закрепление знаний студента, развитие аналитических навыков по проблематике учебной дисциплины, активизацию учебно-познавательной деятельности студентов и снижение аудиторной нагрузки. Часть программного материала выносится для самостоятельного внеаудиторного изучения с последующим текущим или итоговым контролем знаний на занятиях или экзамене. Контроль СРС и оценка ее результатов организуется как самоконтроль (самооценка) студента, а также как контроль и оценка со стороны преподавателя, например в ходе собеседования. Баллы, полученные по СРС студентом, обязательно учитываются при итоговой аттестации по курсу. Формы контроля СРС включают: тестирование; устную беседу по теме с преподавателем; выполнение индивидуального задания и др.

Роль студента в СРС - самостоятельно организовывать свою учебную работу по предложенному преподавателем, методически обеспеченному плану. СРС по курсу учитывает индивидуальные особенности слушателей и включает не только задания, связанные с решением типовых задач, но также творческие задания, требующие самостоятельно «добывать» знания из разных областей, группировать и концентрировать их в контексте конкретной решаемой задачи. Технология обучения предусматривает выработку навыков презентации результатов выполненного индивидуального задания и создание условий для командной работы над комплексной темой с распределением функций и ответственности между членами коллектива. Оценка результатов выполнения индивидуального задания осуществляется по критериям, известным студентам, отражающим наиболее значимые аспекты контроля за выполнением этого вида работ.

<b>Разделы и темы для самостоятельного изучения</b>	<b>Виды и содержание самостоятельной работы</b>
Анализ текстовых данных	-конспектирование первоисточников и другой учебной литературы; -проработка учебного материала (по конспектам лекций учебной и научной литературе) и подготовка докладов на семинарах и практических занятиях, к участию в тематических дискуссиях; -поиск и обзор научных публикаций и электронных источников информации, подготовка заключения по обзору; -работа с тестами и вопросами для самопроверки; - решение домашних контрольных задач.
Нейронные сети	-конспектирование первоисточников и другой учебной литературы; -проработка учебного материала (по конспектам лекций учебной и научной литературе) и подготовка докладов на семинарах и практических занятиях, к участию в тематических дискуссиях; -поиск и обзор научных публикаций и электрон-

	<p>ных источников информации, подготовка заключения по обзору;</p> <p>- работа с тестами и вопросами для самопроверки;</p> <p>- решение домашних контрольных задач.</p>
Обучение с подкреплением. A/B-тестирование и многоармие бандиты	<p>- конспектирование первоисточников и другой учебной литературы;</p> <p>- проработка учебного материала (по конспектам лекций учебной и научной литературе) и подготовка докладов на семинарах и практических занятиях, к участию в тематических дискуссиях;</p> <p>- поиск и обзор научных публикаций и электронных источников информации, подготовка заключения по обзору;</p> <p>- работа с тестами и вопросами для самопроверки;</p> <p>- решение домашних контрольных задач.</p>
Применение методов машинного обучения - ассоциативные правила.	<p>- конспектирование первоисточников и другой учебной литературы;</p> <p>- проработка учебного материала (по конспектам лекций учебной и научной литературе) и подготовка докладов на семинарах и практических занятиях, к участию в тематических дискуссиях;</p> <p>- поиск и обзор научных публикаций и электронных источников информации, подготовка заключения по обзору;</p> <p>- работа с тестами и вопросами для самопроверки;</p> <p>- решение домашних контрольных задач.</p>
Применение методов машинного обучения для анализа социальных сетей	<p>- конспектирование первоисточников и другой учебной литературы;</p> <p>- проработка учебного материала (по конспектам лекций учебной и научной литературе) и подготовка докладов на семинарах и практических занятиях, к участию в тематических дискуссиях;</p> <p>- поиск и обзор научных публикаций и электронных источников информации, подготовка заключения по обзору;</p> <p>- работа с тестами и вопросами для самопроверки;</p> <p>- решение домашних контрольных задач.</p>
Среда Anaconda, Jupyter Notebook. Библиотеки Python: Pandas, NumPy, Matplotlib, SciPy, Документация Python по функциям: optimize, cvxpy, cvxopt.modeling	Изучение документации программного обеспечения

## 7. Фонд оценочных средств для проведения текущего контроля успеваемости, промежуточной аттестации по итогам освоения дисциплины.

### 7.1. Типовые контрольные задания

#### Комплект тестовых заданий

1. Что, из ниже перечисленного, относится к обучающей выборке?

- а) объекты с известными ответами
- б) классификация данных
- в) алгоритм решающий функцию

2. Что называют данными в машинном обучении?

- а) матрицы
- б) объекты
- в) признаки
- г) алгоритм
- д) функция

3. Выберите правильный ответ. Задача классификации - это:

- а) множество объектов, разделенных на классы
- б) исследование влияние одного или нескольких признаков на объект
- в) определение порядка признака согласно рангу

4. Выберите правильный ответ. Задача регрессии - это:

- а) исследование влияние одного или нескольких признаков на объект
- б) множество объектов, разделенных на классы
- в) определение порядка признака согласно рангу

5. Что является традиционной метрикой для задачи регрессии?

- а)  $\sum (y - y_x)^2$
- б)  $\sum |y - y_x|$
- в)  $\sum (y - y_x)$
- г)  $\sum (y - \bar{y})^2$

6. Дана матрица ошибок алгоритма классификации. Вычислите точность прогноза (precision)

	$y = 1$	$y = 0$
$\hat{y} = 1$	15	10
$\hat{y} = 0$	5	20

- а) 0,6
- б) 0,4
- в) 0,7
- г) 0,3

7. Дана матрица ошибок алгоритма классификации. Вычислите долю правильных ответов алгоритма (accuracy):

	$y = 1$	$y = 0$
$\hat{y} = 1$	15	10
$\hat{y} = 0$	5	20

- а) 0,7
- б) 0,4
- в) 0,6
- г) 0,3

8. Дана матрица ошибок алгоритма классификации. Вычислите полноту правильных ответов алгоритма (recall):

	$y = 1$	$y = 0$
$\hat{y} = 1$	15	10

$\hat{y} = 0$	5	20
---------------	---	----

- а) 0,75
- б) 0,67
- в) 0,33
- г) 0,25

9. Предположим, у вас есть 5 ядер свертки в первом слое сверточной нейронной сети. Каждое ядро свертки имеет размер  $7 \times 7$ , имеет нулевое заполнение и имеет шаг 1. Размер входного изображения этого слоя составляет  $224 \times 224 \times 3$ . Каков размер выходного сигнала этого слоя?

- а)  $218 \times 218 \times 5$
- б)  $217 \times 217 \times 3$
- в)  $217 \times 217 \times 8$
- г)  $220 \times 220 \times 7$

10) Какие из следующих форм алгоритм K-Means может не агрегировать?

- а) Спиральное распределение
- б) Круговое распределение
- в) Ленточное распространение
- г) Выпуклое распределение полигонов

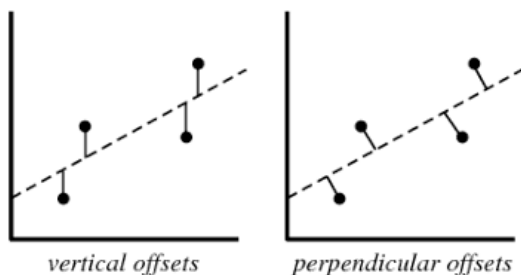
11. При каких условиях метод k-ближайшего соседа работает лучше?

- а) Небольшой образец, но хорошая типичность
- б) Больше образцов, но менее типично
- в) Образцы распределяются по кластерам.
- г) Образцы распределяются по цепочке

12. В целом, какой из следующих методов обычно используется для прогнозирования непрерывных независимых переменных?

- а) Линейная регрессия
- б) Логический обзор
- в) И линейная регрессия и логистическая регрессия
- г) Ничего из перечисленного

13. Какие из следующих смещений мы используем в случае подгонки линии наименьших квадратов? На рисунке абсцисса является входом X, а ордината - выходом Y.



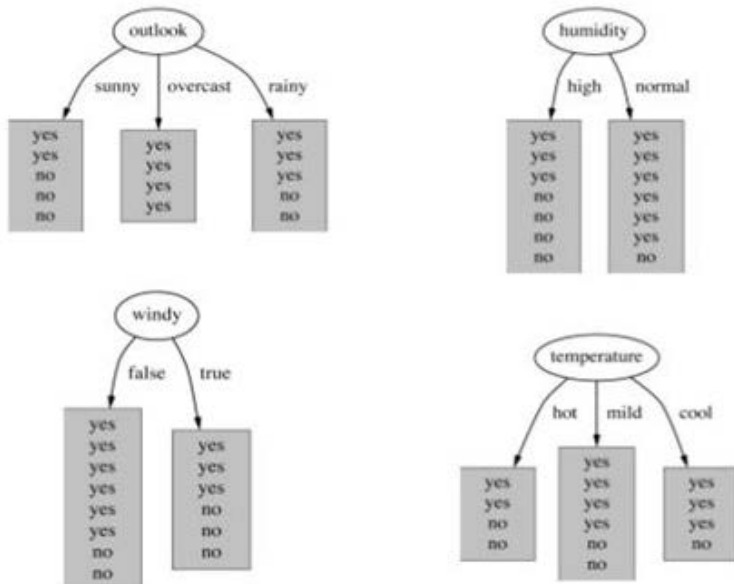
- а) Вертикальные смещения
- б) перпендикулярные смещения
- в) оба смещения в порядке
- г) Ничего из перечисленного

14. В n-мерном пространстве ( $n > 1$ ) какой из следующих методов наиболее подходит для обнаружения выбросов?

- а) расстояние Махаланобиса

- б) Нормальная вероятность графика
- в) Коробочный сюжет
- г) поле корреляции

15. При построении модели дерева решений мы разделяем узел для определенного атрибута. Какое из четырех изображений ниже соответствует наибольшему приросту информации?



- а) outlook
- б) humidity
- в) windy
- г) temperature

16. Линейная регрессия в среде Python реализуется пакетом:  
 а) Scikit-learn б) Pandas в) cvxopt.modeling г) ECOS

### Вопросы для контрольных работ, устного опроса и промежуточного контроля

1. Сущность и этапы анализа данных с помощью методов машинного обучения
2. Классификация методов машинного обучения
3. Основные проблемы машинного обучения
4. Задачи машинного обучения с учителем
5. Задачи машинного обучения без учителя,
6. Задачи машинного обучения с подкреплением.
7. Библиотеки машинного обучения языка Python.
8. Метод регрессии
9. Метод главных компонент.
10. Методы кластеризации в машинном обучении
11. Метод k-средних (k-means) и его модификации
12. Деревья решений, алгоритм случайного леса.
13. Метод опорных векторов.
14. Метод k-ближайших соседей
15. Логистическая регрессия
16. Бинарная линейная классификация
17. Композиции моделей градиентный бустинг и блендинг
18. Нейронные сети, их обучение методом обратного распространения ошибки.

19. Задача кластеризации. Карты Кахоннена.

## **7.2. Методические материалы, определяющие процедуру оценивания знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций.**

### **а) Критерии оценивания компетенций (результатов).**

Программой дисциплины в целях проверки прочности усвоения материала предусматривается проведение различных форм контроля:

1. Текущий контроль – это проверка полноты знаний по основному материалу дисциплинарного модуля (ДМ).

2. Промежуточный контроль - итоговая проверка уровня знаний студента по данной дисциплине в конце семестра (в форме устного или письменного экзамена, сетевого компьютерного тестирования.) Промежуточной формой контроля знаний, умений и навыков по дисциплине является экзамен.

Общий результат выводится как интегральная оценка, складывающаяся из текущего контроля - 50% и промежуточного контроля - 50%.

Текущий контроль по дисциплине включает:

- посещение занятий - 10 баллов,
- выполнение лабораторных заданий – 50 баллов,
- выполнение домашних контрольных работ (самостоятельная работа) - 10 баллов.

Текущий контроль по ДМ:

письменная контрольная работа -15 баллов;

тестирование – 15 баллов;

Промежуточный контроль по дисциплине включает:

- устный экзамен (тестирование) - 100 баллов,

*Критерии оценки посещения занятий* – оценка выставляется по 100 бальной системе и соответствует проценту занятий, которые посетил студент из всего количества аудиторных занятий предусмотренных ДМ.

*Решение задач.*

86-100 баллов выставляется, если студент демонстрирует знания теоретического и практического материала по теме практической работы, определяет взаимосвязи между показателями задачи, даёт правильный алгоритм решения, определяет междисциплинарные связи по условию задания.

66-85 баллов выставляется, если студент демонстрирует знания теоретического и практического материала по теме практической работы, допуская незначительные неточности при решении задач, имея неполное понимание междисциплинарных связей при правильном выборе алгоритма решения задания.

51-65 балл выставляется, если студент затрудняется с правильной оценкой предложенной задачи, даёт неполный ответ, требующий наводящих вопросов преподавателя, выбор алгоритма решения задачи возможен при наводящих вопросах преподавателя.

0-50 баллов выставляется студенту, если он даёт неверную оценку ситуации, неправильно выбирает алгоритм решения.

*Критерии оценки выполнения лабораторных заданий.*

Основными показателями оценки выполненной студентом и представленной для проверки работы являются:

1. Степень соответствия выполненного задания поставленным целям, задачам и требованиям;
2. Оформление, структурирование и комментирование лабораторной работы;
3. Уникальность выполнения работы (отличие от работ коллег);
4. Успешные ответы на контрольные вопросы.

Критерии оценки лабораторной работы.



86-100 баллов - оформление соответствует требованиям, критерии выдержаны, защита всего перечня контрольных вопросов.

66-85 баллов - оформление соответствует требованиям, критерии выдержаны, защита только до 85 % контрольных вопросов.

51-65 балл - оформление соответствует требованиям, критерии выдержаны, защита только до 51 % контрольных вопросов.

0-50 баллов – оформление не соответствует требованиям, критерии не выдержаны, защита только менее 51 % контрольных вопросов.

*Критерии оценки выполнения домашних контрольных работ (самостоятельная работа).*

Основными показателями оценки выполненной студентом и представленной для проверки домашней контрольной работы являются:

1. Степень соответствия выполненного задания поставленным целям, задачам и требованиям;

2. Оформление, структурирование и комментирование лабораторной работы;

3. Уникальность выполнения работы (отличие от работ коллег);

4. Успешные ответы на контрольные вопросы.

Критерии оценки домашней контрольной работы.

86-100 баллов - студент правильно выполнил индивидуальное самостоятельное задание. Показал отличное владения навыками применения полученных знаний и умений при решении профессиональных задач в рамках усвоенного учебного материала. Ответил на все дополнительные вопросы на защите.

66-85 баллов - студент выполнил индивидуальное самостоятельное задание с небольшими неточностями. Показал хорошие владения навыками применения полученных знаний и умений при решении профессиональных задач в рамках усвоенного учебного материала. Ответил на большинство дополнительных вопросов на защите.

51-65 балл - студент выполнил индивидуальное самостоятельное задание с существенными неточностями. Показал удовлетворительное владение навыками применения полученных знаний и умений при решении профессиональных задач в рамках усвоенного учебного материала. При ответах на дополнительные вопросы на защите было допущено много неточностей.

0-50 баллов – при выполнении индивидуального самостоятельного задания студент продемонстрировал недостаточный уровень владения умениями и навыками при решении профессиональных задач в рамках усвоенного учебного материала. При ответах на дополнительные вопросы на защите было допущено множество неточностей.

*Критерии оценки текущего контроля по ДМ (письменная контрольная работа и тестирование).*

*Письменная контрольная работа* состоит из двух типов вопросов:

1. Теоретические вопросы из курса лекций и практических работ. - 40 баллов.

2. Практические вопросы и задачи по лекционному и практическому материалу. - 60 баллов.

86-100 баллов - студент, показал всестороннее, систематическое и глубокое знание учебного материала, самостоятельно ответил на вопросы, ответ отличается богатством и точностью использованных терминов, материал излагается последовательно и логично; показал отличное владения навыками применения полученных знаний и умений при решении профессиональных задач.

66-85 баллов - студент, показал полное знание учебного материала, не допускающий в ответе существенных неточностей, самостоятельно выполнивший ответивший на вопросы; показал хорошие владения навыками применения полученных знаний и умений при решении профессиональных задач

51-65 балл - студент, обнаруживший знание основного учебного материала в объёме, необходимом для дальнейшей учебы самостоятельно выполнивший задания, однако допустивший некоторые погрешности при ответе на вопросы; показал

удовлетворительное владение навыками применения полученных знаний и умений при решении профессиональных задач.

0-50 баллов – выставляется студенту, обнаружившему пробелы в знаниях или отсутствие знаний по значительной части основного учебного материала, не выполнившему задания, допустившему принципиальные ошибки при ответе на вопросы, продемонстрировал недостаточный уровень владения умениями и навыками при решении профессиональных задач.

Критерии выставления оценок за *тестирование* Тестовое задание состоит из пятнадцати вопросов. Время выполнения работы: 15-20 мин.

86-100 баллов - оценка «отлично» – 13-15 правильных ответов;

66-85 баллов - оценка «хорошо» – 10-12 правильных ответов;

51-65 балл - оценка «удовлетворительно» – 8-9 правильных ответов;

0-50 баллов – оценка «неудовлетворительно» – менее 8 правильных ответов.

*Критерии оценки устного экзамена*

Экзаменационные билеты включают три типа заданий:

1. Теоретические вопросы из курса лекций и практических работ. - 30 баллов.

2. Практические вопросы по лекционному и практическому материалу. - 40 баллов.

3. Проблемные вопросы и расчетные задачи. - 30 баллов.

В проверка качества подготовки студентов на экзаменах заканчивается выставлением отметок по принятой пятибалльной шкале («отлично», «хорошо», «удовлетворительно», «неудовлетворительно»).

86-100 баллов - оценка «отлично» - студент владеет знаниями по дисциплине «Методы машинного обучения» в полном объеме учебной программы, достаточно глубоко осмысливает дисциплину; самостоятельно, в логической последовательности и исчерпывающе отвечает на все вопросы билета, подчеркивая при этом самое существенное, умеет анализировать, сравнивать, классифицировать, обобщать, конкретизировать и систематизировать изученный материал, выделять в нем главное; устанавливать причинно-следственные связи; четко формирует ответы, свободно решает ситуационные задачи повышенной сложности; хорошо знаком с основной литературой; увязывает теоретические аспекты дисциплины с прикладными задачами исследования операций и методов оптимизации; владеет современными информационными технологиями решения прикладных задач.

66-85 баллов - оценка «хорошо» – студент владеет знаниями дисциплины «Методы машинного обучения» почти в полном объеме программы (имеются пробелы знаний только в некоторых, особенно сложных разделах); самостоятельно и отчасти при наводящих вопросах дает полноценные ответы на вопросы билета; не всегда выделяет наиболее существенное, не допускает вместе с тем серьезных ошибок в ответах; умеет решать легкие и средней тяжести ситуационные задачи; умеет трактовать выбор тех или иных методов и средств решения прикладных задач.

51-65 балл - оценка «удовлетворительно» - студент владеет основным объемом знаний по дисциплине «Методы машинного обучения»; проявляет затруднения в самостоятельных ответах, оперирует неточными формулировками; в процессе ответов допускаются ошибки по существу вопросов. Студент способен решать лишь наиболее легкие задачи, владеет только обязательным минимумом при решении задач исследования операций.

0-50 баллов – оценка «неудовлетворительно» - студент не освоил обязательного минимума знаний дисциплины «Методы машинного обучения», не способен ответить на вопросы билета даже при дополнительных наводящих вопросах экзаменатора.

*Критерии оценки экзамена в форме тестирования*

Тестовое задание состоит из тридцати вопросов. Время выполнения работы: 60 мин.

86-100 баллов - оценка «отлично» – 26-30 правильных ответов;

66-85 баллов - оценка «хорошо» – 20-25 правильных ответов;

51-65 балл - оценка «удовлетворительно» – 16-19 правильных ответов;  
0-50 баллов – оценка «неудовлетворительно» – менее 16 правильных ответов.

## 8. Учебно-методическое обеспечение дисциплины.

### а) адрес сайта курса

<http://eor.dgu.ru/>.

### б) основная литература:

1. Неделько, В. М. Основы статистических методов машинного обучения: учебное пособие / В. М. Неделько. - Новосибирск: Новосибирский государственный технический университет, 2010. - 72 с. - ISBN 978-5-7782-1385-2. - Текст: электронный // Электронно-библиотечная система IPR BOOKS: [сайт]. - URL: <https://www.iprbookshop.ru/45418.html> (дата обращения: 11.02.2022)
2. Павлова, А. И. Информационные технологии: основные положения теории искусственных нейронных сетей: учебное пособие / А. И. Павлова. - Новосибирск : Новосибирский государственный университет экономики и управления «НИИХ», 2017. - 191 с. - ISBN 978-5-7014-0801-0. - Текст: электронный // Электронно-библиотечная система IPR BOOKS: [сайт]. - URL: <https://www.iprbookshop.ru/87110.html> (дата обращения: 11.02.2022)
3. Теория и практика машинного обучения: учебное пособие / В. В. Воронина, А. В. Михеев, Н. Г. Ярушкина, К. В. Святков. - Ульяновск: Ульяновский государственный технический университет, 2017. - 291 с. - ISBN 978-5-9795-1712-4. - Текст: электронный // Электронно-библиотечная система IPR BOOKS: [сайт]. - URL: <https://www.iprbookshop.ru/106120.html> (дата обращения: 11.02.2022)

### в) дополнительная литература:

1. Ракитский, А. А. Методы машинного обучения: учебно-методическое пособие / А. А. Ракитский. - Новосибирск: Сибирский государственный университет телекоммуникаций и информатики, 2018. - 32 с. - Текст: электронный // Электронно-библиотечная система IPR BOOKS: [сайт]. - URL: <https://www.iprbookshop.ru/90591.html> (дата обращения: 22.06.2022).
2. Сараев, П. В. Методы машинного обучения: методические указания и задания к лабораторным работам по курсу / П. В. Сараев. - Липецк: Липецкий государственный технический университет, ЭБС АСВ, 2017. - 48 с. - Текст: электронный // Электронно-библиотечная система IPR BOOKS: [сайт]. - URL: <https://www.iprbookshop.ru/83183.html>. (дата обращения: 22.02.2022).
3. Anaconda. Individual Edition [Электронный ресурс]. - Режим доступа: URL: <https://www.anaconda.com/> (дата обращения: 22.02.2022).
4. Pandas - Python Data Analysis Library [Электронный ресурс]. - Режим доступа: URL: <https://pandas.pydata.org/> (дата обращения: 22.02.2022).
5. Rashi Desai Топ-10 библиотек Python для Data Science. [Электронный ресурс]. - Режим доступа: - Режим доступа: URL: <https://towardsdatascience.com/top-10-python-libraries-for-data-science-cd82294ec266> (дата обращения: 22.02.2022).
6. Scikit-learn. Машинное обучение на Python. - Режим доступа: URL: <https://scikit-learn.org/>

## 9. Перечень ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины.

1. eLIBRARY.RU [Электронный ресурс]: электронная библиотека. – Режим доступа: <http://elibrary.ru/defaultx.asp> (дата обращения: 22.02.2022). – Яз. рус., англ.

2. IPRbooks [Электронный ресурс]: Электронно-библиотечная система. - Режим доступа: <http://www.iprbookshop.ru/> (дата обращения: 11.02.2020). – Яз. рус., англ.
3. Университетская библиотека онлайн [Электронный ресурс]: Электронно-библиотечная система. - Режим доступа: <http://biblioclub.ru/> (дата обращения: 22.02.2022). – Яз. рус., англ.
4. Электронный каталог НБ ДГУ [Электронный ресурс]: база данных содержит сведения о всех видах лит, поступающих в фонд НБ ДГУ/Дагестанский гос. ун-т. – Махачкала, 2010 – Режим доступа: <http://elib.dgu.ru>, свободный (дата обращения: 22.02.2022).

#### **10. Методические указания для обучающихся по освоению дисциплины.**

Для изучения теоретического курса студентам необходимо использовать лекционный материал, учебники и учебные пособия из списка основной и дополнительной литературы, интернет источники.

По дисциплине «Методы машинного обучения» в конце каждого модуля проводится контрольная работа.

В контрольную работу включаются теоретические вопросы и задачи тех типов, которые были разобраны на предшествующих практических занятиях.

Рабочей программой дисциплины «Методы машинного обучения» предусмотрена самостоятельная работа студентов в объеме 152 (170) часов. Самостоятельная работа проводится с целью углубления знаний по дисциплине и предусматривает:

- чтение студентами рекомендованной литературы и усвоение теоретического материала дисциплины;

- выполнение индивидуальных заданий;

- подготовку к контрольным работам, зачету и экзаменам.

С самого начала изучения дисциплины студент должен четко уяснить, что без систематической самостоятельной работы успех невозможен. Эта работа должна регулярно начинаться сразу после лекционных и практических занятий, для закрепления только что пройденного материала.

После усвоения теоретического материала можно приступить к самостоятельному решению задач из учебников и пособий, входящих в список основной литературы.

#### **11. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине, включая перечень программного обеспечения и информационных справочных систем.**

Microsoft Office (Excel, Power Point), Anaconda, Python, Jupyter Notebook

#### **12. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине.**

Компьютерный класс, аудитория для проведения лекционных и практических занятий и самостоятельной работы средствами оборудованная оргтехникой, персональными компьютерами, объединенными в сеть с выходом в Интернет; установленное лицензионное и свободное программное обеспечение.