

**МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«ИНГУШСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»**

**ФИЗИКО- МАТЕМАТИЧЕСКИЙ ФАКУЛЬТЕТ
КАФЕДРА «ИНФОРМАЦИОННЫЕ СИСТЕМЫ И ТЕХНОЛОГИИ»**

СОГЛАСОВАНО

УТВЕРЖДАЮ

Руководитель образовательной программы

И.о декана физико-математического
факультета

_____/М.Х. Мальсагов
«20» мая 2024г.

_____/Б.С.Кульбужев
«23» мая 2024г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

Б1.В.13 Машинное обучение

Направление подготовки (бакалавриат)

09.03.02 «Информационные системы и технологии»

Направленность (профиль подготовки)

Технологии искусственного интеллекта и анализа данных

Квалификация выпускника

Бакалавр

Форма обучения

Очная, очно-заочная

Магас, 2024 г.

Рабочая программа дисциплины «Машинное обучение» составлена в соответствии с требованиями ФГОСВО по направлению подготовки 09.03.02-«Информационные системы и технологии», утвержденного приказом Министерства образования и науки Российской Федерации от «19» сентября 2017 г. № 926.

Программу составила: старший преподаватель кафедры «Информационные системы и технологии», _____/Евлоева З.Д.

Программа одобрена на заседании кафедры «Информационные системы и технологии»

Протокол № 9 от «20» мая 2024 года

Программа одобрена Учебно-методической комиссией физико-математического факультета

Протокол № 9 от « 22 » мая 2024 года

1. ЦЕЛИ И ЗАДАЧИ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Цель дисциплины – познакомить студентов с научными и инженерными проблемами интеллектуального обеспечения вычислительных процессов, методами, средствами разработки и эксплуатации систем машинного обучения. Кроме того, преследуется цель обеспечить понимание студентами принципов глубокого обучения, а также отладки интеллектуальных алгоритмов.

Задачами дисциплины являются:

В результате освоения дисциплины студент должен:

Знать:

- основные модели знаний и характеристики систем искусственного интеллекта, методы и средства разработки интеллектуальных систем и баз знаний;

- основы методов машинного обучения;

- методы организации и основные процессы интеллектуального обслуживания вычислительных процессов;

Уметь:

- определять интеллектуальные потребности вычислительных процессов, выполнять интеллектуальное обслуживание вычислительных систем; разрабатывать и применять средства интеллектуальной поддержки вычислительных процессов.

- отлаживать интеллектуальные алгоритмы. Приобрести навыки:

- определения интеллектуальных потребностей вычислительных процессов; выполнения интеллектуального обслуживания вычислительных систем;

- разработки и применения средств интеллектуальной поддержки вычислительных процессов.

Формируемые дисциплиной знания и умения готовят выпускника данной образовательной программы к выполнению следующих обобщенных трудовых функций (трудовых функций):

Код и наименование профессионального стандарта	Обобщенные трудовые функции			Трудовые функции		
	Код	Наименование	Уровень квалификации	Наименование	Код	Уровень (подуровень) квалификации
06.015 Специалист по информационным системам.	С	Выполнение работ и управление работами по созданию (модификации) и сопровождению ИС, автоматизирующих задачи	6	Определение первоначальных требований заказчика к ИС и возможности их реализации в ИС на этапе пред контрактных работ	С/01.6	6
				Документирование существующих бизнес-процессов организации заказчика (реверс-инжиниринг бизнес- процессов организации)	С/07.6	6

		организационного управления и бизнес-процессы.	Разработка модели бизнес-процессов заказчика	C/08.6	6
			Разработка архитектуры ИС	C/14.6	6
			Проектирование и дизайн ИС	C/16.6	6
			Разработка баз данных ИС	C/17.6	6

2. Место учебной дисциплины в структуре основной профессиональной образовательной программы

Учебная дисциплина «Машинное обучение» относится к модулю Б1.В.13 обязательной программы, формируемой участниками образовательных отношений. В соответствии с учебным планом дисциплина изучается в четвертом семестре.

Связь дисциплины «Машинное обучение» с предшествующими дисциплинами и сроки их изучения

Код дисциплины	Дисциплины, предшествующие дисциплине «Машинное обучение»	Семестр
Б1.0.07	«Информатика»	1
Б1.0.19	«Математика»	1,2
Б1.0.12	«Технологии программирования»	3
Б1.0.16	«Базы данных»	3

Связь дисциплины «Машинное обучение» с последующими дисциплинами и сроки их изучения

Код дисциплины	Дисциплины, следующие за дисциплиной «Машинное обучение»	Семестр
Б1.В.05	Программирование промышленных логических контроллеров	7,8
Б1.0.15	Лингвистическое обеспечение робототехнических систем	6

Связь дисциплины «Машинное обучение» со смежными дисциплинами и сроки их изучения

Код дисциплины	Дисциплины, смежные с дисциплиной «Машинное обучение»	Семестр
Б1.В.16	Технология анализа и визуализация данных	5

3. Результаты освоения дисциплины «Машинное обучение»

Освоение дисциплины направлено на формирование следующих компетенций в соответствии с образовательным стандартом и образовательной программой:

Перечень планируемых результатов обучения по дисциплине,

соотнесённых с индикаторами достижения компетенций

Категория (группа) общепрофессиональных компетенций	Код и наименование общепрофессионально й компетенции	Код и наименование индикатора достижения общепрофессиональной компетенции
ПК-4	ПК-4 Способен применять методы машинного обучения для решения задач искусственного интеллекта	<p>ИД-1 ПК-4 Принимает участие в оценке, выборе и при необходимости разработке методов машинного обучения:</p> <p>Принимает участие в оценке, выборе и при необходимости разработке методов машинного обучения:</p> <ul style="list-style-type: none"> - Знает: классические методы и алгоритмы машинного обучения: предиктивные - обучение с учителем, дескриптивные - обучение без учителя; - Умеет: проводить сравнительный анализ и осуществлять выбор, настройку при необходимости разработку методов и алгоритмов для решения задач машинного обучения.
ПК-5	ПК-5Способен использовать инструментальные средства для решения задач машинного обучения.	<p>ИД-1 ПК-5 Знать: Осуществлять оценку и выбор инструментальных средств для решения поставленной задачи;</p> <p>ИД-2 ПК-5 Уметь: Разрабатывать модели машинного обучения для решения задач;</p> <p>ИД-3 ПК-5 Владеть: Создает, поддерживает и использует системы искусственного интеллекта, включающие разработанные модели и методы, с применением выбранных инструментов машинного обучения:</p>
ПК-6	ПК-6 Способен создавать и поддерживать системы искусственного интеллекта на основе нейросетевых моделей и методов..	<p>ИД-1 ПК-6 Осуществляет оценку и выбор моделей искусственных нейронных сетей и инструментальных средств для решения поставленной задачи:</p> <ul style="list-style-type: none"> - Знает базовые архитектуры и модели искусственных нейронных сетей; - Знает функциональность современных инструментальных средств и систем программирования в области создания моделей искусственных нейронных сетей; - Умеет проводить оценку и выбор моделей искусственных нейронных сетей и инструментальных средств для решения задачи машинного обучения; <p>Владеет: применять современные инструментальные средства и системы программирования для разработки и обучения моделей искусственных нейронных сетей</p>

ПК-7.	ПК-7 Способен осуществлять сбор и подготовку данных для систем искусственного интеллекта.	ИД-1 ПК-7 Выполняет подготовку и разметку структурированных и неструктурированных данных для машинного обучения. - Знает: методы редукции размерности элементов набора данных и их предварительной статистической обработки, разметки структурированных и неструктурированных данных; - Умеет: выявлять и исключать из массива данных ошибочные данные и выбросы; - Умеет выделять входные и выходные переменные с целью использования предиктивных моделей; осуществлять разметку структурированных и неструктурированных данных; - Владеет: использовать инструменты, библиотеки и технологии Data Science для подготовки и разметки структурированных и неструктурированных данных для машинного обучения; использовать методы и технологии массово параллельной обработки и анализа данных.
-------	-------------------------------------------------------------------------------------------	-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

4. Структура и содержание дисциплины (модуля)

«Машинное обучение»

4.1 Структура дисциплины (модуля)

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

№ п/п	Наименование разделов и тем дисциплины (модуля)	семестр	Виды учебной работы, включая самостоятельную работу студентов трудоемкость (в часах)								Формы текущего контроля успеваемости (по неделям семестра) Форма промежуточной аттестации (по семестрам)									
			Контактная работа					Самостоятельная работа												
			Всего	Лекции	Практические занятия	Лабораторные занятия	Др. виды контакт. работы	Всего	Курсовая работа(проект)	Подготовка к экзамену	Другие виды СР	Собеседование	Коллоквиум	Проверка тестов	Проверка контрол.н. работ	Проверка реферата	Проверка эссе и иных творческих работ	курсовая работа (проект) др.		
1	Раздел 1. Основы машинного обучения и методы регрессивного анализа	5	6	2	4	-	-	-	-	-	-	-	6	-	-	-	-	-	-	-
2	Раздел 2. Методы классификации	5	6	2	4	-	-	-	-	-	-	-	8	-	-	-	-	-	-	-

3	Раздел 3. Методы обучения без учителя	5	6	2	4	-	-	-	-	6	-	-	-	-	-	-	-
4	Раздел 4. Нейронные сети и глубокое обучение	5	10	4	6	-	-	-	-	8	-	-	-	-	-	-	-
5	Раздел 5. Отладка программ машинного обучения.	5	6	2	4	-	-	-	-	8	-	-	-	-	-	-	-
6	Раздел 6. Деревья решений	5	6	2	4	-	-	-	-	6	-	-	-	-	-	-	-
	Раздел 7. Ансамблевые методы	5	10	4	6	-	-	-	-	8	-	-	-	-	-	-	-
	Подготовка к промежуточной аттестации, консультации перед промежуточной аттестацией и сдача аттестации.	5	-	-	-	-	-	-	-	8	-	-	-	-	-	-	-
	Зачет	5	9	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	Общая трудоемкость, в часах		50	18	32	-	-	-	-	58	-	-	-	-	-	-	-

4.2.Содержание учебного материала

Тема 1. Основы машинного обучения и методы регрессионного анализа

Введение в машинное обучение. Проблема линейной регрессии. Функция ошибки. Градиентный спуск

Тема 2. Методы классификации

Проблема классификации. Сигмоидальная функция. Функция ошибки. Логистическая регрессия

Деревья решений, случайный лес. Метод ближайших соседей

Тема 3. Методы обучения без учителя

Методы снижения размерностей (PCA, T-SNE). Обзор методов кластеризации. K-Means

Агломеративная кластеризация. Иерархические методы кластеризации. Метрики оценки качества кластеризации

Тема 4. Нейронные сети и глубокое обучение

Сложные проблемы классификации. Аппроксимация сложных функций. Прямое распространение

Обратное распространение. Сверточные сети. LSTM сети. GAN сети

Тема 5. Отладка программ машинного обучения

Проблема переобучения и недообучения. Отладка алгоритмов машинного обучения

Тема 6. Деревья решений

Структура деревьев решений. Виды разделяющих функций. Обучения дерева

решений. Алгоритм Random Forest.

Тема 7. Ансамблевые методы

Ансамблевые методы Ансамбли алгоритмов машинного обучения. Агрегирование моделей. Ансамбли решающих деревьев. Метод случайного леса. Градиентный бустинг

Перечень практических работ

№ п/ п	Название практических работ	Количество часов
1	П.р. № 1. Основы работы с данными в python.	8
2	П.р. № 2. Линейная регрессия, полиномиальные функции, регуляризация	4
3	П.р. № 3. Логистическая регрессия, нейронные сети	4
4	П.р. № 4. Кластеризация	4
5	П.р. № 5. Глубокое обучение	4
6	П.р. № 6. Правила и анализ качества (точность, полнота). Анализ с помощью ROC кривой. Алгоритм построения деревьев решений.	4
7	П.р. № 7. Ансамблевые методы Реализация моделей с помощью метода градиентного бустинга, метода случайного леса. Блендинг и стеккинг	4
Всего часов		32

5. Образовательные технологии, включая интерактивные формы обучения

В организации процесса обучения по дисциплине используются как традиционные технологии, характерные для лекционно-семинарской формы обучения, так и инновационные технологии. Эффективность обучения повышает использование современных технических средств и методик изучения предмета.

Лекционная аудитория с мультимедиа проектором, компьютером, стандартным набором специализированной учебной мебели и учебного оборудования, персональные компьютеры. На каждом персональном компьютере обеспечен выход в сеть Internet, установлен пакет офисных программ.

6. Учебно-методическое обеспечение самостоятельной работы студентов.

Оценочные средства для текущего контроля успеваемости, промежуточной аттестации по итогам освоения дисциплины.

Перечень самостоятельных работ

№ п/ п	Название самостоятельной работы	Количество часов
1	С. № 1. Проблема линейной регрессии. Функция ошибки. Градиентный спуск	6
2	№ 2 Проблема классификации. Сигмоидальная функция. Логистическая регрессия	8

3	№ 3. Методы снижения размерностей (PCA, T-SNE). Обзор методов кластеризации. K-Means. . Иерархические методы кластеризации. Метрики оценки качества кластеризации	6
4	№4.Сложные проблемы классификации. Обратное распространение. Сверточные сети. LSTM сети. GAN сети	8
5	№ 5 Проблема переобучения и недообучения. Отладка алгоритмов машинного обучения	8
6	№ 6. Критерий информационного выигрыша и критерий Джини. Леса решающих деревьев.	6
7	Реализация моделей с помощью метода градиентного бустинга, метода случайного леса. Методы отбора признаков. Оптимизация гиперпараметров	8
8	Подготовка к промежуточной аттестации, консультации перед промежуточной аттестацией и сдача аттестации.	8
Всего часов		58

6.1. Методические указания по организации самостоятельной работы студентов

1. Успешное освоение курса требует напряженной самостоятельной работы студента. В программе курса приведено минимально необходимое время для работы студента над темой. Самостоятельная работа включает в себя чтение лекций и рекомендованной литературы, решение задач, предлагаемых студентам на лекциях и лабораторных занятиях, разбор проблемных ситуаций. Руководство и контроль за самостоятельной работой студента осуществляется в форме индивидуальных консультаций. Для активизации самостоятельной работы студентов и экономии времени, отводимого на лекционный курс, ряд тем выносятся на самостоятельное изучение. Самостоятельная работа со студентами проводится в часы самостоятельной работы в форме консультаций. Распределение часов руководства самостоятельной работой учитывает важность рассматриваемой темы и возможную сложность при освоении ее студентами. Самостоятельная работа студентов рассматривается как вид учебного труда, позволяющий целенаправленно формировать и развивать самостоятельность студента как личностное качество при выполнении различных видов заданий и проработке дополнительного учебного материала. Для успешного выполнения лабораторных работ, написания рефератов и подготовки к коллоквиуму, помимо материалов лекционных и лабораторных занятий, необходимо использовать основную и дополнительную литературу, указанную в конце данной рабочей программы.

2. Лекции, презентации, методические указания и задания к лабораторным работам помещаются в групповые папки студентов, находящиеся на сервере университета и доступны студентам группы.

3. Методические указания содержат теорию по рассматриваемому вопросу, рекомендации по выполнению практических работ.

6.2. Оценочные средства для текущего контроля успеваемости, промежуточной аттестации по итогам освоения дисциплины

Итоговый контроль проводится в виде зачета.

Контрольные вопросы

1. Препроцессинг. Масштабирование. Нормировка. Полиномиальные признаки. One-hot encoding.
2. Типы обучения: с учителем, без учителя, с подкреплением, с частичным участием учителя, активное обучение.
3. Смещение и дисперсия (bias and variance). Понятие средней гипотезы.
4. Ошибка внутри и вне выборки. Ошибка обобщения. Неравенство Хёфдинга. Валидация и кросс-валидация.
5. Линейная регрессия.
6. Полиномиальная регрессия.
7. Гребневая регрессия.
8. Логистическая регрессия. Градиентный спуск.
9. Байесовский классификатор. Типы оценки распределений признаков (Gaussian, Bernoulli, Multinomial). EM алгоритм.
10. Метод опорных векторов. Постановка задачи. Формулировка и решение двойственной задачи. Типы опорных векторов. Ядра.
11. Пороговые условия. Эффективность по Парето.
12. Precision-Recall и ROC кривые. AUC.
13. Метрические классификаторы. kNN. WkNN. Отбор эталонов. DROP5. Kdtree.
14. Кластеризация. kMeans, MeanShift, DBSCAN, Affinity Propagation.
15. Деревья решений. Информационный выигрыш, критерий Джини.
16. Регуляризация деревьев. Небрежные решающие деревья.
17. Бустинг деревьев решений.
18. Ансамблевые методы. Soft and Hard Voting. Bagging.
19. Случайные леса. AdaBoost.
20. Ансамблевые методы регрессии. RANSAC. Theil-Sen. Huber.

7. Учебно-методическое и материально-техническое обеспечение дисциплины (модуля) «Машинное обучение»

7.1 Основная литература

1. Воронина, В. В. Теория и практика машинного обучения : учебное пособие / В. В. Воронина. — Ульяновск : УлГТУ, 2017. — 290 с. — ISBN 978-5-9795-1712-4. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/165053> (дата обращения: 20.12.2021). — Режим доступа: для авториз. пользователей.
2. Ростовцев, В. С. Искусственные нейронные сети : учебник для вузов / В. С. Ростовцев. — 2-е изд., стер. — Санкт-Петербург : Лань, 2021. — 216 с. — ISBN 978-5-8114-7462-2. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/160142> (дата обращения: 20.12.2021). — Режим доступа: для авториз. пользователей.
3. Евдокимов А.А. Системное программирование [Электронный ресурс. Мультимедиа]: учебное пособие / А. А. Евдокимов, Н. В. Майстренко, А. В. Майстренко. - Тамбов: ФГБОУ ВО "ТГТУ", 2016. - Режим доступа к книге: <http://www.tstu.ru/book/elib3/mm/2016/evdokimov>
4. Павлова, А. И. Информационные технологии: основные положения теории искусственных нейронных сетей : учебное пособие / А. И. Павлова.

- Новосибирск : Новосибирский государственный университет экономики и управления «НИНХ», 2017. — 191 с. — ISBN 978-5-7014-0801-0. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <http://www.iprbookshop.ru/87110.html> (дата обращения: 14.01.2021). — Режим доступа: для авторизир. пользователей
5. Барский, А. Б. Введение в нейронные сети : учебное пособие / А. Б. Барский. — 3-е изд. — Москва, Саратов : Интернет-Университет Информационных Технологий (ИНТУИТ), Ай Пи Ар Медиа, 2020. — 357 с. — ISBN 978-5-4497-0309-5. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <http://www.iprbookshop.ru/89426.html> (дата обращения: 14.01.2021). — Режим доступа: для авторизир. пользователей

7.2 Дополнительная литература

1. <http://learn.fist.ulstu.ru/> - Курс «Машинное обучение»
2. Барский, А. Б. Искусственный интеллект и логические нейронные сети : учебное пособие / А. Б. Барский. — Санкт-Петербург : Интермедия, 2019. — 360 с. — ISBN 978-5-4383-0155-4. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <http://www.iprbookshop.ru/95270.html> (дата обращения: 14.01.2021). — Режим доступа: для авторизир. пользователей

7.3. Используемые ресурсы информационно-телекоммуникационной сети «Internet»; информационные технологии, программные средства и информационно-справочные системы

Наряду с традиционными изданиями студенты и сотрудники имеют возможность пользоваться электронными полнотекстовыми базами данных:

Название ресурса	Ссылка/доступ
Электронная библиотека онлайн «Единое окно к образовательным ресурсам»	http://window.edu.ru
«Образовательный ресурс России»	http://school-collection.edu.ru
Федеральный образовательный портал: учреждения, программы, стандарты, ВУЗы, тесты ЕГЭ, ГИА	http://www.edu.ru
Федеральный центр информационно-образовательных ресурсов (ФЦИОР)	http://fcior.edu.ru
Русская виртуальная библиотека	http://rvb.ru
Кабинет русского языка и литературы	http://ruslit.ioso.ru
Национальный корпус русского языка	http://ruscorpora.ru
Научная электронная библиотека «e-Library»	http://elibrary.ru/defaultx.asp
Электронно-библиотечная система IPRbooks	http://www.iprbookshop.ru
Электронно-библиотечная система ИнГГУ	https://lib.inggu.ru/
Информационно-правовая система «Гарант»	Сетевая версия, доступна со всех компьютеров в корпоративной сети ИнГГУ

Информационно-библиотечное обеспечение учебного процесса включает в себя:

- доступ к электронно-библиотечным системам и электронным документам;
 - хранение выпускных работ и ведения электронного портфолио обучающихся;
 - WV-reader (IPRbooks) для мобильных устройств для незрячих и слабовидящих.
- Имеющиеся в вузе адаптивные технологии для внедрения инклюзивного

образования обеспечивают возможность внедрения методов инклюзивного образования для обучения людей с нарушениями зрения в формах, адаптированных к ограничениям их здоровья.

7.4. Программное обеспечение

Университет обеспечен необходимым комплектом лицензионного и свободно распространяемого программного обеспечения, в том числе отечественного производства.

Каждый обучающийся в течение всего периода обучения обеспечен индивидуальным неограниченным доступом к электронной информационно-образовательной среде университета из любой точки, в которой имеется доступ к информационно-телекоммуникационной сети «Интернет» как на территории университета, так и вне ее.

Лицензионное программное обеспечение:

1. Microsoft Windows 7, Windows 8, Windows 8.1, Windows 10
2. Microsoft Windows server 2003, 2008, 2012, 2016
3. MicrosoftOffice 2007, 2010, 2016
4. Антивирусное ПО Kaspersky endpoint security
5. Справочно-правовая система «Гарант»

7.5. Материально-техническое обеспечение дисциплины

Материально-техническая база университета позволяет обеспечивать качественное проведение теоретических и практических занятий.

Для проведения лекций по дисциплине используются специализированные аудитории с мультимедийным оборудованием или с возможностями подключения к такому оборудованию, позволяющему демонстрировать на большом экране приемы работы с персональным компьютером и другой лекционный материал (технические характеристики компьютера, входящего в состав мультимедийного оборудования или используемого совместно с таким оборудованием, должны обеспечивать возможность работы с современными версиями операционной системы Windows, пакета Microsoft Office, обслуживающих, программы Corel Draw и Adobe Photoshop, прикладных программ и другого, в том числе и сетевого программного обеспечения).

Для проведения лабораторных занятий по дисциплине и для самостоятельной работы студентов используются специализированные аудитории, оснащенные терминалами и персональными компьютерами, подключенными к центральному серверу, обеспечивающему технические характеристики обслуживания терминалов или персональных компьютеров, позволяющие при проведении лабораторных занятий использовать современное программное обеспечение (операционную систему Windows 7 и выше, пакет Microsoft Office 2010 и выше, программы Gimp, Corel Draw, а также обслуживающие программы и среды разработки программ по выбору преподавателей).