

Министерство науки и высшего образования Российской Федерации  
Федеральное государственное автономное образовательное учреждение  
высшего образования  
«Уральский федеральный университет имени первого Президента России Б.Н. Ельцина»

УТВЕРЖДАЮ  
Директор по образовательной  
деятельности

\_\_\_\_\_ С.Т. Князев  
«\_\_» \_\_\_\_\_

### РАБОЧАЯ ПРОГРАММА МОДУЛЯ

Код модуля	Модуль
1163488	Основы машинного обучения

Екатеринбург

<b>Перечень сведений о рабочей программе модуля</b>	<b>Учетные данные</b>
<b>Образовательная программа</b> 1. Биотехнические системы и технологии	<b>Код ОП</b> 1. 12.03.04/33.01
<b>Направление подготовки</b> 1. Биотехнические системы и технологии	<b>Код направления и уровня подготовки</b> 1. 12.03.04

Программа модуля составлена авторами:

<b>№ п/п</b>	<b>Фамилия Имя Отчество</b>	<b>Ученая степень, ученое звание</b>	<b>Должность</b>	<b>Подразделение</b>
1	Смирнов Андрей Алексеевич	кандидат физико-математических наук, без ученого звания	Доцент	УрФУ. Кафедра Экспериментальной физики

**Согласовано:**

Управление образовательных программ

Р.Х. Токарева

# 1. ОБЩАЯ ХАРАКТЕРИСТИКА МОДУЛЯ **Основы машинного обучения**

## 1.1. Аннотация содержания модуля

Изучение модуля позволит студентам познакомиться с моделями и методами интеллектуального анализа данных и машинного обучения в задачах обработки и анализа данных, а также приобрести навыки исследователя данных (data scientist) и разработчика математических моделей, методов и алгоритмов анализа данных. В рамках освоения модуля студенты изучают основные модели (линейные, метрические, логические), подходы к их обучению и методы обработки данных.

## 1.2. Структура и объем модуля

Таблица 1

№ п/п	Перечень дисциплин модуля в последовательности их освоения	Объем дисциплин модуля и всего модуля в зачетных единицах
1	Основы машинного обучения	3
ИТОГО по модулю:		3

## 1.3. Последовательность освоения модуля в образовательной программе

Пререквизиты модуля	
Постреквизиты и кореквизиты модуля	1. Анализ и обработка биомедицинских данных

## 1.4. Распределение компетенций по дисциплинам модуля, планируемые результаты обучения (индикаторы) по модулю

Таблица 2

Перечень дисциплин модуля	Код и наименование компетенции	Планируемые результаты обучения (индикаторы)
1	2	3
Основы машинного обучения	ПК-3 - Способен проводить экспериментальные исследования и измерения, обрабатывать и представлять полученные данные с учетом специфики	З-1 - Излагать основные приемы и методы проведения экспериментальных исследований и изысканий, которые могут быть использованы для решения поставленных прикладных задач, относящихся к области биотехнических систем и технологий У-2 - Определять перечень необходимых ресурсов и временные затраты при

	биотехнических систем и технологий	<p>составлении плана проведения исследований и изысканий</p> <p>П-2 - Составлять план проведения экспериментальных исследований и изысканий, включающий перечень необходимых ресурсов и временные затраты</p>
	<p>ПК-4 - Способен использовать современные информационные технологии и программное обеспечение при решении задач профессиональной деятельности, соблюдая требования информационной безопасности</p>	<p>З-3 - Объяснять принципы создания информации в цифровой форме и ее использование в информационных процессах</p> <p>З-5 - Приводить примеры применения информационных сервисов для решения поставленных задач</p> <p>У-1 - Формулировать особенности использования информационных технологий и программного обеспечения в предметной области биотехнических систем и технологий</p> <p>У-3 - Выбирать конфигурацию вычислительной системы, операционную систему, пакеты прикладных программ, информационные сервисы и базы данных для обработки, передачи и хранения информации в цифровой форме</p> <p>П-1 - Выполнять поставленные задачи по поиску, обработке, передаче и хранению информации в цифровой форме, используя современные технические средства, пакеты прикладных программ, информационные сервисы и базы данных</p> <p>П-2 - Обосновывать выбор технических и программных средств защиты персональных данных и данных организации при работе с информационными системами на основе анализа потенциальных и реальных угроз безопасности информации</p>

### 1.5. Форма обучения

Обучение по дисциплинам модуля может осуществляться в очной формах.

**РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ**  
**Основы машинного обучения**

Рабочая программа дисциплины составлена авторами:

<b>№ п/п</b>	<b>Фамилия Имя Отчество</b>	<b>Ученая степень, ученое звание</b>	<b>Должность</b>	<b>Подразделение</b>
1	Смирнов Андрей Алексеевич	кандидат физико- математических наук, без ученого звания	Доцент	экспериментально й физики

**Рекомендовано учебно-методическим советом института Физико-технологический**

Протокол № 7 от 15.03.2024 г.

# 1. СОДЕРЖАНИЕ И ОСОБЕННОСТИ РЕАЛИЗАЦИИ ДИСЦИПЛИНЫ

Авторы:

- Смирнов Андрей Алексеевич, Доцент, экспериментальной физики

## 1.1. Технологии реализации, используемые при изучении дисциплины модуля

- Традиционная (репродуктивная) технология
- Разноуровневое (дифференцированное) обучение
  - Базовый уровень

*\*Базовый I уровень – сохраняет логику самой науки и позволяет получить упрощенное, но верное и полное представление о предмете дисциплины, требует знание системы понятий, умение решать проблемные ситуации. Освоение данного уровня результатов обучения должно обеспечить формирование запланированных компетенций и позволит обучающемуся на минимальном уровне самостоятельности и ответственности выполнять задания;*

*Продвинутый II уровень – углубляет и обогащает базовый уровень как по содержанию, так и по глубине проработки материала дисциплины. Это происходит за счет включения дополнительной информации. Данный уровень требует умения решать проблемы в рамках курса и смежных курсов посредством самостоятельной постановки цели и выбора программы действий. Освоение данного уровня результатов обучения позволит обучающемуся повысить уровень самостоятельности и ответственности до творческого применения знаний и умений.*

## 1.2. Содержание дисциплины

Таблица 1.1

Код раздела, темы	Раздел, тема дисциплины*	Содержание
1	Основные понятия и задачи в машинном обучении	Что такое машинное обучение? Основные определения. Постановки задач. Виды признаков. Загрузка данных, базовый анализ. Работа с таблицей на примере удаления столбца. Пропущенные значения, описание данных. Значения признаков, выбор ячеек с помощью условий. Извлечение новых признаков, распределение значений, среднее. Группировка и агрегирование.
2	Метод k ближайших соседей	Сравнение объектов и метрики. Измерение ошибки модели. Оценка обобщающей способности. Метод k ближайших соседей с весами. Метод k ближайших соседей для регрессии. Поиск оптимальных гиперпараметров. Классификация с методом ближайших соседей.
3	Линейная регрессия	Модель линейной регрессии. Применение линейной регрессии. Линейная регрессия в векторном виде. Обучение линейной регрессии. Переобучение и регуляризация. Интерпретация линейных моделей.
4	Градиентный спуск	Градиент и его свойства. Градиентный спуск. Стохастический градиентный спуск. Функции потерь. Вывод градиента для MSE. Векторное дифференцирование. Регрессия в матричной форме. Градиентный спуск в матричной форме. Реализация градиентного спуска. Настройка градиентного спуска.
5	Методы линейной классификации	Линейный классификатор. Обучение линейных классификаторов. Метрики качества классификации. Совмещение точности и полноты. Метрики качества

		ранжирования. Площадь под ROC-кривой. Обучение линейного классификатора с помощью градиентного спуска. Вероятности классов, AUC-PRC и AUC-ROC. Логистическая регрессия. Метод опорных векторов. Калибровка вероятностей. Многоклассовая классификация.
6	Решающие деревья	Нелинейные модели. Решающие деревья. Структура решающих деревьев. Выбор предикатов. Критерии информативности. Жадное построение дерева. Сравнение линейных моделей и решающих деревьев. Решающие деревья - переобучение и неустойчивость. Реализации критерия ошибки. Визуализация зависимости критерия ошибки от порога разбиения. Поиск лучшего разбиения и интерпретация. Влияние масштабирования признаков на решающее дерево.
7	Композиции: бэггинг, блендинг и стэкинг	Неустойчивость деревьев. Композиции моделей. Смещение и разброс моделей. Случайный лес. Регрессия решающего дерева. Важность признаков.
8	Градиентный бустинг	Идея бустинга. Бустинг для MSE. Произвольная функция потерь для бустинга. Градиентный бустинг в общем виде. Гиперпараметры и регуляризация. Имплементации бустинга. Работа с категориальными признаками. Сравнение градиентного бустинга и случайного леса. Свойства градиентного бустинга. Современные библиотеки градиентного бустинга.
9	Обучение без учителя	Кластеризация. Обучение без учителя. Плотностный алгоритм для кластеризации пространственных данных с присутствием шума (DBSCAN). Метод главных компонент (PCA). Стохастическое вложение соседей с t-распределением (t-SNE). K-means на практике. Кластеризация на реальных данных. Понижение размерности с помощью PCA. Понижение размерности с помощью t-SNE.

### 1.3. Направление, виды воспитательной деятельности и используемые технологии

Таблица 1.2

Направление воспитательной деятельности	Вид воспитательной деятельности	Технология воспитательной деятельности	Компетенция	Результаты обучения
Профессиональное воспитание	целенаправленная работа с информацией для использования в практических целях	Технология анализа образовательных задач	ПК-4 - Способен использовать современные информационные технологии и программное обеспечение при решении задач профессиональной деятельности, соблюдая требования	П-1 - Выполнять поставленные задачи по поиску, обработке, передаче и хранению информации в цифровой форме, используя современные технические средства, пакеты

			информационной безопасности	прикладных программ, информационные сервисы и базы данных
--	--	--	-----------------------------	---

1.4. Программа дисциплины реализуется на государственном языке Российской Федерации .

## **2. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ**

### **Основы машинного обучения**

#### **Электронные ресурсы (издания)**

1. Неделько, В. М.; Основы статистических методов машинного обучения : учебное пособие.; Новосибирский государственный технический университет, Новосибирск; 2010; <http://www.iprbookshop.ru/45418.html> (Электронное издание)
2. Кухаренко, Б. Г.; Интеллектуальные системы и технологии : учебное пособие.; Московская государственная академия водного транспорта, Москва; 2015; <http://www.iprbookshop.ru/47933.html> (Электронное издание)
3. Сараев, П. В.; Методы машинного обучения : методические указания и задания к лабораторным работам по курсу.; Липецкий государственный технический университет, ЭБС АСВ, Липецк; 2017; <http://www.iprbookshop.ru/83183.html> (Электронное издание)

#### **Печатные издания**

1. Лимановская, О. В., Обабков, И. Н.; Основы машинного обучения : учебное пособие для студентов вуза, обучающихся по направлениям подготовки: 09.03.04 - Программная инженерия; 09.03.03 - Прикладная информатика; 02.04.02 - Фундаментальная информатика и информационные технологии; 09.04.02 - Информационные системы и технологии.; Издательство Уральского университета, Екатеринбург; 2020 (15 экз.)

#### **Профессиональные базы данных, информационно-справочные системы**

1. Полнотекстовая БД American Physical Society (<https://journals.aps.org/about>).
2. Полнотекстовая БД eLibrary - научная электронная библиотека (<http://elibrary.ru>).
3. Полнотекстовая БД Institute of Physics (IOP) (<http://iopscience.iop.org/>).

#### **Материалы для лиц с ОВЗ**

Весь контент ЭБС представлен в виде файлов специального формата для воспроизведения синтезатором речи, а также в тестовом виде, пригодном для прочтения с использованием экранной лупы и настройкой контрастности.

#### **Базы данных, информационно-справочные и поисковые системы**

1. Информационная база данных по биомедицинской инженерии, режим доступа <http://www.physionet.org>
2. Видеоportal по медико-биологическим вопросам, режим доступа: <http://www.med-edu.ru>



3. Государственная публичная научно-техническая библиотека, режим доступа: <http://www.gpntb.ru>

4. Электронная библиотека нормативно-технической документации, режим доступа: <http://www.technormativ.ru>

5. Зональная научная библиотека УрФУ, режим доступа: <http://lib.urfu.ru>

### 3. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

#### Основы машинного обучения

#### Сведения об оснащённости дисциплины специализированным и лабораторным оборудованием и программным обеспечением

Таблица 3.1

№ п/п	Виды занятий	Оснащённость специальных помещений и помещений для самостоятельной работы	Перечень лицензионного программного обеспечения
1	Лекции	Мебель аудиторная с количеством рабочих мест в соответствии с количеством студентов Рабочее место преподавателя Доска аудиторная Периферийное устройство	Office Professional 2003 Win32 Russian CD-ROM
2	Практические занятия	Мебель аудиторная с количеством рабочих мест в соответствии с количеством студентов Рабочее место преподавателя Доска аудиторная	Не требуется
3	Лабораторные занятия	Мебель аудиторная с количеством рабочих мест в соответствии с количеством студентов Рабочее место преподавателя Персональные компьютеры по количеству обучающихся	Office Professional 2003 Win32 Russian CD-ROM
4	Консультации	Мебель аудиторная с количеством рабочих мест в соответствии с количеством студентов Рабочее место преподавателя	Не требуется

5	Текущий контроль и промежуточная аттестация	<p>Мебель аудиторная с количеством рабочих мест в соответствии с количеством студентов</p> <p>Рабочее место преподавателя</p> <p>Подключение к сети Интернет</p>	Office Professional 2003 Win32 Russian CD-ROM
6	Самостоятельная работа студентов	<p>Мебель аудиторная с количеством рабочих мест в соответствии с количеством студентов</p> <p>Персональные компьютеры по количеству обучающихся</p> <p>Подключение к сети Интернет</p>	Office Professional 2003 Win32 Russian CD-ROM