

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ  
Федеральное государственное автономное учреждение  
высшего профессионального образования  
"Казанский (Приволжский) федеральный университет"  
Институт вычислительной математики и информационных технологий



**УТВЕРЖДАЮ**

Проректор  
по образовательной деятельности КФУ  
Проф. Таюрский Д.А.

\_\_\_\_\_ 20\_\_ г.

**Программа дисциплины**  
Машинное обучение Б1.В.ДВ.9

Направление подготовки: 01.03.02 - Прикладная математика и информатика

Профиль подготовки: Системное программирование

Квалификация выпускника: бакалавр

Форма обучения: очное

Язык обучения: русский

**Автор(ы):**

Разинков Е.В.

**Рецензент(ы):**

Столов Е.Л.

**СОГЛАСОВАНО:**

Заведующий(ая) кафедрой: Латыпов Р. Х.

Протокол заседания кафедры No \_\_\_\_ от " \_\_\_\_ " \_\_\_\_\_ 201\_\_ г

Учебно-методическая комиссия Института вычислительной математики и информационных технологий:

Протокол заседания УМК No \_\_\_\_ от " \_\_\_\_ " \_\_\_\_\_ 201\_\_ г

Регистрационный No

Казань  
2016

## Содержание

1. Цели освоения дисциплины
2. Место дисциплины в структуре основной образовательной программы
3. Компетенции обучающегося, формируемые в результате освоения дисциплины /модуля
4. Структура и содержание дисциплины/ модуля
5. Образовательные технологии, включая интерактивные формы обучения
6. Оценочные средства для текущего контроля успеваемости, промежуточной аттестации по итогам освоения дисциплины и учебно-методическое обеспечение самостоятельной работы студентов
7. Литература
8. Интернет-ресурсы
9. Материально-техническое обеспечение дисциплины/модуля согласно утвержденному учебному плану

Программу дисциплины разработал(а)(и) ассистент, к.н. Разинков Е.В. кафедра системного анализа и информационных технологий отделение фундаментальной информатики и информационных технологий , Evgenij.Razinkov@kpfu.ru

### 1. Цели освоения дисциплины

Цель освоения дисциплины заключается в ознакомлении с базовыми понятиями машинного обучения, с основными алгоритмами машинного обучения, особенностями их применения. Данная область является алгоритмической основой многих современных информационных технологий и имеет множество прикладных применений.

### 2. Место дисциплины в структуре основной образовательной программы высшего профессионального образования

Данная учебная дисциплина включена в раздел " Б1.В.ДВ.9 Дисциплины (модули)" основной образовательной программы 01.03.02 Прикладная математика и информатика и относится к дисциплинам по выбору. Осваивается на 4 курсе, 7 семестр.

Для освоения дисциплины студентам необходимы знания из линейной алгебры, теории вероятностей, математического анализа. Полученные в процессе изучения дисциплины знания могут использоваться при написании курсовых и дипломных работ, при изучении алгоритмов компьютерного зрения и интеллектуального анализа данных.

### 3. Компетенции обучающегося, формируемые в результате освоения дисциплины /модуля

В результате освоения дисциплины формируются следующие компетенции:

Шифр компетенции	Расшифровка приобретаемой компетенции
ОПК-1 (профессиональные компетенции)	способностью использовать базовые знания естественных наук, математики и информатики, основные факты, концепции, принципы теорий, связанных с прикладной математикой и информатикой
ОПК-2 (профессиональные компетенции)	способностью приобретать новые научные и профессиональные знания, используя современные образовательные и информационные технологии
ОПК-4 (профессиональные компетенции)	способностью решать стандартные задачи профессиональной деятельности на основе информационной и библиографической культуры с применением информационно-коммуникационных технологий и с учетом основных требований информационной безопасности

В результате освоения дисциплины студент:

1. должен знать:

- возможности алгоритмов машинного обучения;
- классы задач, решаемых с помощью алгоритмов машинного обучения.

2. должен уметь:

- применять на практике алгоритмы машинного обучения;
- обосновать применение того или иного алгоритма машинного обучения для решения конкретной задачи.

3. должен владеть:

- базовым инструментарием машинного обучения;

- программно реализовывать алгоритмы машинного обучения;
- применять алгоритмы машинного обучения на практике;
- анализировать результаты обучения алгоритма, предлагать пути повышения точности алгоритма.

#### 4. Структура и содержание дисциплины/ модуля

Общая трудоемкость дисциплины составляет 2 зачетных(ые) единиц(ы) 72 часа(ов).

Форма промежуточного контроля дисциплины зачет в 7 семестре.

Суммарно по дисциплине можно получить 100 баллов, из них текущая работа оценивается в 50 баллов, итоговая форма контроля - в 50 баллов. Минимальное количество для допуска к зачету 28 баллов.

86 баллов и более - "отлично" (отл.);

71-85 баллов - "хорошо" (хор.);

55-70 баллов - "удовлетворительно" (удов.);

54 балла и менее - "неудовлетворительно" (неуд.).

#### 4.1 Структура и содержание аудиторной работы по дисциплине/ модулю

##### Тематический план дисциплины/модуля

N	Раздел Дисциплины/ Модуля	Семестр	Неделя семестра	Виды и часы аудиторной работы, их трудоемкость (в часах)			Текущие формы контроля
				Лекции	Практические занятия	Лабораторные работы	
1.	Тема 1. Введение в машинное обучение	7		0	4	0	письменное домашнее задание
2.	Тема 2. Линейные модели регрессии	7		0	4	0	письменное домашнее задание
3.	Тема 3. Логистическая регрессия	7		0	4	0	письменное домашнее задание
4.	Тема 4. Нейронные сети	7		0	6	0	контрольная работа письменное домашнее задание
5.	Тема 5. Деревья решений	7		0	6	0	письменное домашнее задание
7.	Тема 7. Алгоритм AdaBoost	7		0	6	0	письменное домашнее задание

N	Раздел Дисциплины/ Модуля	Семестр	Неделя семестра	Виды и часы аудиторной работы, их трудоемкость (в часах)			Текущие формы контроля
				Лекции	Практические занятия	Лабораторные работы	
8.	Тема 8. Кластеризация	7		0	6	0	контрольная работа
	Тема . Итоговая форма контроля	7		0	0	0	зачет
	Итого			0	36	0	

#### 4.2 Содержание дисциплины

##### Тема 1. Введение в машинное обучение

###### *практическое занятие (4 часа(ов)):*

Обзор задач, решаемых алгоритмами машинного обучения. Классификация алгоритмов машинного обучения.

##### Тема 2. Линейные модели регрессии

###### *практическое занятие (4 часа(ов)):*

Линейная регрессия. Линейные модели регрессии. Базисные функции. Регуляризация. Программная реализация алгоритма линейной регрессии.

##### Тема 3. Логистическая регрессия

###### *практическое занятие (4 часа(ов)):*

Целевая функция логистической регрессии. Регуляризация логистической регрессии. Программная реализация алгоритма логистической регрессии.

##### Тема 4. Нейронные сети

###### *практическое занятие (6 часа(ов)):*

Структура нейрона. Структура нейронной сети. Обучения нейронной сети с помощью алгоритма обратного распространения ошибки. Программная реализация алгоритма обратного распространения ошибки.

##### Тема 5. Деревья решений

###### *практическое занятие (6 часа(ов)):*

Структура деревьев решений. Виды разделяющих функций. Обучения дерева решений. Алгоритм Random Forest. Программная реализация дерева решений.

##### Тема 7. Алгоритм AdaBoost

###### *практическое занятие (6 часа(ов)):*

Описание алгоритма AdaBoost. Математическое обоснование алгоритма. Каскад классификаторов. Программная реализация алгоритма AdaBoost.

##### Тема 8. Кластеризация

###### *практическое занятие (6 часа(ов)):*

Обзор существующих алгоритмов классификации. Алгоритм k-means. Программная реализация алгоритма k-means.

#### 4.3 Структура и содержание самостоятельной работы дисциплины (модуля)

N	Раздел Дисциплины	Семестр	Неделя семестра	Виды самостоятельной работы студентов	Трудоемкость (в часах)	Формы контроля самостоятельной работы
---	----------------------	---------	--------------------	--	---------------------------	---

1.	Тема 1. Введение в					
----	--------------------	--	--	--	--	--

машинное обучение

7

Изучение  
литературы по

приложениям машинного обучения.

задание



N	Раздел Дисциплины	Семестр	Неделя семестра	Виды самостоятельной работы студентов	Трудоемкость (в часах)	Формы контроля самостоятельной работы
2.	Тема 2. Линейные модели регрессии	7		Проведение экспериментов.	4	домашнее задание
3.	Тема 3. Логистическая регрессия	7		Проведение экспериментов.	4	домашнее задание
4.	Тема 4. Нейронные сети	7		подготовка к контрольной работе	4	контрольная работа
				Проведение экспериментов.	4	домашнее задание
5.	Тема 5. Деревья решений	7		подготовка домашнего задания	4	домашнее задание
7.	Тема 7. Алгоритм AdaBoost	7		подготовка домашнего задания	4	домашнее задание
8.	Тема 8. Кластеризация	7		подготовка к контрольной работе	8	контрольная работа
	Итого				36	

## 5. Образовательные технологии, включая интерактивные формы обучения

Обучение происходит в форме лабораторных занятий, а также самостоятельной работы студентов.

Теоретический материал излагается на аудиторных занятиях лекционного типа, а также предполагает изучение студентами самостоятельно. Изучение курса подразумевает не только овладение теоретическим материалом, но и получение практических навыков для более глубокого понимания разделов на основе примеров и упражнений, иллюстрирующих рассматриваемые теоретические положения. Аудиторные занятия подразумевают интерактивные обсуждения различных примеров или в активной форме самостоятельного выполнения заданий студентами.

Самостоятельная работа предполагает выполнение домашних заданий. Практические задания, выполненные в аудитории, предназначены для указания общих методов решения задач определенного типа. Закрепить навыки можно лишь в результате самостоятельной работы.

Кроме того, самостоятельная работа включает подготовку к зачету. При подготовке к сдаче зачета весь объем работы рекомендуется распределять равномерно по дням, отведенным для подготовки к зачету, контролировать каждый день выполнения работы.

## 6. Оценочные средства для текущего контроля успеваемости, промежуточной аттестации по итогам освоения дисциплины и учебно-методическое обеспечение самостоятельной работы студентов

### Тема 1. Введение в машинное обучение

домашнее задание , примерные вопросы:

Разбор возможных приложений алгоритмов машинного обучения.

## **Тема 2. Линейные модели регрессии**

домашнее задание , примерные вопросы:

Проведение экспериментов с программной реализацией линейной регрессии. Исследование влияния параметров алгоритма на значение целевой функции на обучающей и тестовой выборке.

## **Тема 3. Логистическая регрессия**

домашнее задание , примерные вопросы:

Проведение экспериментов с программной реализацией логистической регрессии. Исследование влияния параметров алгоритма на точность алгоритма на обучающей и тестовой выборке.

## **Тема 4. Нейронные сети**

домашнее задание , примерные вопросы:

Проведение экспериментов с программной реализацией нейронной сети. Исследование влияния параметров алгоритма на точность алгоритма на обучающей и тестовой выборке.

контрольная работа , примерные вопросы:

Программная реализация нейронной сети для различных задач. Использование различных разбиений множества примеров на обучающую и тестовую выборки.

## **Тема 5. Деревья решений**

домашнее задание , примерные вопросы:

Проведение экспериментов с программной реализацией дерева решений. Исследование влияния параметров алгоритма на точность алгоритма на обучающей и тестовой выборке.

## **Тема 7. Алгоритм AdaBoost**

домашнее задание , примерные вопросы:

Проведение экспериментов с программной реализацией алгоритма AdaBoost. Исследование влияния параметров слабых классификаторов на точность алгоритма на обучающей и тестовой выборке.

## **Тема 8. Кластеризация**

контрольная работа , примерные вопросы:

Программная реализация каскада с деревьями решений в качестве слабых классификаторов.

## **Тема . Итоговая форма контроля**

Примерные вопросы к зачету:

Вопросы к зачету:

- Классификация алгоритмов машинного обучения.
- Регрессия, линейная регрессия, линейные модели регрессии.
- Целевая функция линейной регрессии.
- Классификация. Логистическая регрессия.
- Целевая функция логистической регрессии.
- Регуляризация.
- Нейронные сети. Структура нейронной сети.
- Алгоритм обратного распространения ошибки.
- Деревья решений. Обучение дерева решений.
- Алгоритм Random Forest.
- Алгоритм AdaBoost.
- Каскад классификаторов. Обучение каскада.
- Кластеризация. Алгоритм k-means.

## 7.1. Основная литература:

1. Джонс, М. Т. Программирование искусственного интеллекта в приложениях [Электронный ресурс] / М. Тим Джонс ; Пер. с англ. А. И. Осипов. - 2-е изд. - М. : ДМК Пресс, 2011. - 312 с.: ил. - ISBN 978-5-94074-746-8.

<http://znanium.com/bookread2.php?book=410211>

2. Саттон, Р. С. Обучение с подкреплением [Электронный ресурс] / Р. С. Саттон, Э. Г. Барто ; пер. с англ. - Эл. изд. - М.: БИНОМ. Лаборатория знаний, 2012. - 399 с.: ил. - (Адаптивные и интеллектуальные системы). - ISBN 978-5-9963-1349-5.

<http://e.lanbook.com/view/book/4405/>

3. Шапиро, Л. Компьютерное зрение [Электронный ресурс] / Л. Шапиро, Дж. Стокман; пер. с англ. - 2-е изд. (эл.). - М.: БИНОМ. Лаборатория знаний, 2013. - 752 с.: ил. - (Лучший зарубежный учебник). - ISBN 978-5-9963-1312-9.

<http://e.lanbook.com/view/book/42631/>

4. Вьюгин, В.В. Математические основы машинного обучения и прогнозирования [Электронный ресурс] - Электрон. дан. - М. : МЦНМО (Московский центр непрерывного математического образования), 2013. - 304 с. - Режим доступа:

[http://e.lanbook.com/books/element.php?pl1\\_id=56397](http://e.lanbook.com/books/element.php?pl1_id=56397)

5. Ясницкий, Л.Н. Искусственный интеллект. Элективный курс : учебное пособие [Электронный ресурс] : учебно-методическое пособие. - Электрон. дан. - М. : "Лаборатория знаний" (ранее "БИНОМ. Лаборатория знаний"), 2012. - 201 с. - Режим доступа:

[http://e.lanbook.com/books/element.php?pl1\\_id=8775](http://e.lanbook.com/books/element.php?pl1_id=8775)

## 7.2. Дополнительная литература:

1. Информационные аналитические системы [Электронный ресурс] : учебник / Т. В. Алексеева, Ю. В. Амириди, В. В. Дик и др.; под ред. В. В. Дика. - М.: МФПУ Синергия, 2013. - 384 с. Режим доступа: <http://znanium.com/bookread2.php?book=451186>

2. Куприянов М. С., Барсегян, А. А. Технологии анализа данных: Data Mining, Visual Mining, Text Mining, OLAP [Электронный ресурс] / А. А. Барсегян, М. С. Куприянов, В. В. Степаненко. - 2-е изд., перераб. и доп. - СПб.: БХВ-Петербург, 2007. - 384 с. Режим доступа:<http://znanium.com/bookread2.php?book=489445>

3. Осипов Г. В. Методы искусственного интеллекта/Осипов Г.В. - М.: Физматлит, 2011. - 296 с. Режим доступа: <http://znanium.com/bookread2.php?book=544787>

4. Введение в специальность программиста: Учебник / В.А. Гвоздева. - 2-е изд., испр. и доп. - М.: ИД ФОРУМ: НИЦ Инфра-М, 2013. - 208 с. Режим доступа:

<http://znanium.com/bookread2.php?book=398911>

## 7.3. Интернет-ресурсы:

Видео-лекции курса "Машинное обучение" от Yandex -

<https://yandexdataschool.ru/edu-process/courses/machine-learning>

Научный портал по математическим наукам - <http://www.mathnet.ru>

Портал образовательных ресурсов по ИТ - <http://www.intuit.ru>

Портал статей по применению ИТ и машинному обучению -

[http://habrahabr.ru/hub/machine\\_learning/](http://habrahabr.ru/hub/machine_learning/)

Профессиональный интернет-ресурс по машинному обучению - <http://www.machinelearning.ru/>

## 8. Материально-техническое обеспечение дисциплины(модуля)

Освоение дисциплины "Машинное обучение" предполагает использование следующего материально-технического обеспечения:

Мультимедийная аудитория, вместимостью более 60 человек. Мультимедийная аудитория состоит из интегрированных инженерных систем с единой системой управления, оснащенная современными средствами воспроизведения и визуализации любой видео и аудио информации, получения и передачи электронных документов. Типовая комплектация мультимедийной аудитории состоит из: мультимедийного проектора, автоматизированного проекционного экрана, акустической системы, а также интерактивной трибуны преподавателя, включающей тач-скрин монитор с диагональю не менее 22 дюймов, персональный компьютер (с техническими характеристиками не ниже Intel Core i3-2100, DDR3 4096Mb, 500Gb), конференц-микрофон, беспроводной микрофон, блок управления оборудованием, интерфейсы подключения: USB, audio, HDMI. Интерактивная трибуна преподавателя является ключевым элементом управления, объединяющим все устройства в единую систему, и служит полноценным рабочим местом преподавателя. Преподаватель имеет возможность легко управлять всей системой, не отходя от трибуны, что позволяет проводить лекции, практические занятия, презентации, вебинары, конференции и другие виды аудиторной нагрузки обучающихся в удобной и доступной для них форме с применением современных интерактивных средств обучения, в том числе с использованием в процессе обучения всех корпоративных ресурсов. Мультимедийная аудитория также оснащена широкополосным доступом в сеть интернет. Компьютерное оборудование имеет соответствующее лицензионное программное обеспечение.

Компьютерный класс, представляющий собой рабочее место преподавателя и не менее 15 рабочих мест студентов, включающих компьютерный стол, стул, персональный компьютер, лицензионное программное обеспечение. Каждый компьютер имеет широкополосный доступ в сеть Интернет. Все компьютеры подключены к корпоративной компьютерной сети КФУ и находятся в едином домене.

Учебно-методическая литература для данной дисциплины имеется в наличии в электронно-библиотечной системе "ZNANIUM.COM", доступ к которой предоставлен студентам. ЭБС "ZNANIUM.COM" содержит произведения крупнейших российских учёных, руководителей государственных органов, преподавателей ведущих вузов страны, высококвалифицированных специалистов в различных сферах бизнеса. Фонд библиотеки сформирован с учетом всех изменений образовательных стандартов и включает учебники, учебные пособия, УМК, монографии, авторефераты, диссертации, энциклопедии, словари и справочники, законодательно-нормативные документы, специальные периодические издания и издания, выпускаемые издательствами вузов. В настоящее время ЭБС ZNANIUM.COM соответствует всем требованиям федеральных государственных образовательных стандартов высшего профессионального образования (ФГОС ВПО) нового поколения.

Учебно-методическая литература для данной дисциплины имеется в наличии в электронно-библиотечной системе Издательства "Лань", доступ к которой предоставлен студентам. ЭБС Издательства "Лань" включает в себя электронные версии книг издательства "Лань" и других ведущих издательств учебной литературы, а также электронные версии периодических изданий по естественным, техническим и гуманитарным наукам. ЭБС Издательства "Лань" обеспечивает доступ к научной, учебной литературе и научным периодическим изданиям по максимальному количеству профильных направлений с соблюдением всех авторских и смежных прав.

Аудитория для проведения лекций с мультимедийным оборудованием, компьютерный класс.

Программа составлена в соответствии с требованиями ФГОС ВПО и учебным планом по направлению 01.03.02 "Прикладная математика и информатика" и профилю подготовки Системное программирование .

Автор(ы):

Разинков Е.В. \_\_\_\_\_

"\_\_" \_\_\_\_\_ 201\_\_ г.

Рецензент(ы):

Столов Е.Л. \_\_\_\_\_

"\_\_" \_\_\_\_\_ 201\_\_ г.