

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.

подписано электронно-цифровой подписью

Программа дисциплины
Машинное обучение Б1.В.ДВ.10

Направление подготовки: 02.03.02 - Фундаментальная информатика и информационные технологии

Профиль подготовки: Системный анализ и информационные технологии

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 96618

Казань
2018

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.

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

Данная учебная дисциплина включена в раздел " Б1.В.ДВ.10 Дисциплины (модули)" основной образовательной программы 02.03.02 Фундаментальная информатика и информационные технологии и относится к дисциплинам по выбору. Осваивается на 3 курсе, 6 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 фундаментальной информатикой и информационными технологиями
ОПК-2 (профессиональные компетенции)	способностью применять в профессиональной деятельности современные языки программирования и языки баз данных, методологии системной инженерии, системы автоматизации проектирования, электронные библиотеки и коллекции, сетевые технологии, библиотеки и пакеты программ, современные профессиональные стандарты информационных технологий
ОПК-3 (профессиональные компетенции)	способностью приобретать новые научные и профессиональные знания, используя современные образовательные и информационные технологии
ОПК-3 (профессиональные компетенции)	способностью к разработке алгоритмических и программных решений в области системного и прикладного программирования, математических, информационных и имитационных моделей, созданию информационных ресурсов глобальных сетей, образовательного контента, прикладных баз данных, тестов и средств тестирования систем и средств на соответствие стандартам и исходным требованиям
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

Шифр компетенции	Расшифровка приобретаемой компетенции
ПК-1 (профессиональные компетенции)	способностью собирать, обрабатывать и интерпретировать данные современных научных исследований, необходимые для формирования выводов по соответствующим научным исследованиям
ПК-4 (профессиональные компетенции)	способностью решать задачи профессиональной деятельности в составе научно-исследовательского и производственного коллектива
ПК-7 (профессиональные компетенции)	способностью разрабатывать и реализовывать процессы жизненного цикла информационных систем, программного обеспечения, сервисов систем информационных технологий, а также методы и механизмы оценки и анализа функционирования средств и систем информационных технологий
ПК-8 (профессиональные компетенции)	способностью применять на практике международные и профессиональные стандарты информационных технологий, современные парадигмы и методологии, инструментальные и вычислительные средства

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;
4. должен демонстрировать способность и готовность:
 - 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 6 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	6	1-2	2	0	2	Письменное домашнее задание
2.	Тема 2. Линейные модели регрессии	6	3-4	2	0	2	Письменное домашнее задание
3.	Тема 3. Логистическая регрессия	6	5-6	2	0	2	Письменное домашнее задание
4.	Тема 4. Нейронные сети	6	7-9	3	0	3	Контрольная работа Письменное домашнее задание
5.	Тема 5. Деревья решений	6	10-12	3	0	3	Письменное домашнее задание
7.	Тема 7. Алгоритм AdaBoost	6	13-15	3	0	3	Письменное домашнее задание
8.	Тема 8. Кластеризация	6	16-18	3	0	3	Контрольная работа
	Тема . Итоговая форма контроля	6		0	0	0	Зачет
	Итого			18	0	18	

4.2 Содержание дисциплины

Тема 1. Введение в машинное обучение

лекционное занятие (2 часа(ов)):

Обзор задач, решаемых алгоритмами машинного обучения.

лабораторная работа (2 часа(ов)):

Классификация алгоритмов машинного обучения.

Тема 2. Линейные модели регрессии

лекционное занятие (2 часа(ов)):

Линейная регрессия. Линейные модели регрессии. Базисные функции. Регуляризация.

лабораторная работа (2 часа(ов)):

Программная реализация алгоритма линейной регрессии.

Тема 3. Логистическая регрессия

лекционное занятие (2 часа(ов)):

Целевая функция логистической регрессии. Регуляризация логистической регрессии.

лабораторная работа (2 часа(ов)):

Программная реализация алгоритма логистической регрессии.

Тема 4. Нейронные сети

лекционное занятие (3 часа(ов)):

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

лабораторная работа (3 часа(ов)):

Программная реализация алгоритма обратного распространения ошибки.

Тема 5. Деревья решений

лекционное занятие (3 часа(ов)):

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

лабораторная работа (3 часа(ов)):

Программная реализация дерева решений.

Тема 7. Алгоритм AdaBoost

лекционное занятие (3 часа(ов)):

Описание алгоритма AdaBoost. Математическое обоснование алгоритма. Каскад классификаторов.

лабораторная работа (3 часа(ов)):

Программная реализация алгоритма AdaBoost.

Тема 8. Кластеризация

лекционное занятие (3 часа(ов)):

Обзор существующих алгоритмов классификации. Алгоритм k-means.

лабораторная работа (3 часа(ов)):

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	6	1-2	Изучение литературы по приложениям машинного обучения.	4	домашнее задание
2.	Тема 2. Линейные модели регрессии	6	3-4	Проведение экспериментов.	4	домашнее задание
3.	Тема 3. Логистическая регрессия	6	5-6	Проведение экспериментов.	4	домашнее задание
4.	Тема 4. Нейронные сети	6	7-9	подготовка к контрольной работе	4	контрольная работа
				Проведение экспериментов.	4	домашнее задание
5.	Тема 5. Деревья решений	6	10-12	подготовка домашнего задания	4	домашнее задание

N	Раздел Дисциплины	Семестр	Неделя семестра	Виды самостоятельной работы студентов	Трудоемкость (в часах)	Формы контроля самостоятельной работы
7.	Тема 7. Алгоритм AdaBoost	6	13-15	подготовка домашнего задания	4	домашнее задание
8.	Тема 8. Кластеризация	6	16-18	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. Шапиро, Л. Компьютерное зрение [Электронный ресурс] / Л. Шапиро, Д. Стокман. ? Электрон. дан. ? Москва : Издательство 'Лаборатория знаний', 2015. ? 763 с. ? Режим доступа: <https://e.lanbook.com/book/84096>.
2. Жданов, А.А. Автономный искусственный интеллект [Электронный ресурс] : учеб. пособие ? Электрон. дан. ? Москва : Издательство 'Лаборатория знаний', 2015. ? 362 с. ? Режим доступа: <https://e.lanbook.com/book/70761>.
3. 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

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. Смолин, Д.В. Введение в искусственный интеллект: конспект лекций [Электронный ресурс] : учеб. пособие ? Электрон. дан. ? Москва : Физматлит, 2007. ? 264 с. ? Режим доступа: <https://e.lanbook.com/book/2325>.
3. Осипов Г. В. Методы искусственного интеллекта/ОсиповГ.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/
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 02.03.02 "Фундаментальная информатика и информационные технологии" и профилю подготовки Системный анализ и информационные технологии .

Автор(ы):

Разинков Е.В. _____

"__" _____ 201__ г.

Рецензент(ы):

Столов Е.Л. _____

"__" _____ 201__ г.