

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное автономное образовательное учреждение высшего образования
"Казанский (Приволжский) федеральный университет"
Институт математики и механики им. Н.И. Лобачевского



УТВЕРЖДАЮ

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

"__" _____ 20__ г.

Программа дисциплины

Машинное обучение

Направление подготовки: 02.04.01 - Математика и компьютерные науки

Профиль подготовки: Методы математического и алгоритмического моделирования общенаучных и прикладных задач

Квалификация выпускника: магистр

Форма обучения: очное

Язык обучения: русский

Год начала обучения по образовательной программе: 2018

Содержание

1. Перечень планируемых результатов обучения по дисциплине (модулю), соотнесенных с планируемыми результатами освоения ОПОП ВО
2. Место дисциплины (модуля) в структуре ОПОП ВО
3. Объем дисциплины (модуля) в зачетных единицах с указанием количества часов, выделенных на контактную работу обучающихся с преподавателем (по видам учебных занятий) и на самостоятельную работу обучающихся
4. Содержание дисциплины (модуля), структурированное по темам (разделам) с указанием отведенного на них количества академических часов и видов учебных занятий
 - 4.1. Структура и тематический план контактной и самостоятельной работы по дисциплине (модулю)
 - 4.2. Содержание дисциплины (модуля)
5. Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине (модулю)
6. Фонд оценочных средств по дисциплине (модулю)
7. Перечень литературы, необходимой для освоения дисциплины (модуля)
8. Перечень ресурсов информационно-телекоммуникационной сети "Интернет", необходимых для освоения дисциплины (модуля)
9. Методические указания для обучающихся по освоению дисциплины (модуля)
10. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине (модулю), включая перечень программного обеспечения и информационных справочных систем (при необходимости)
11. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине (модулю)
12. Средства адаптации преподавания дисциплины (модуля) к потребностям обучающихся инвалидов и лиц с ограниченными возможностями здоровья
13. Приложение №1. Фонд оценочных средств
14. Приложение №2. Перечень литературы, необходимой для освоения дисциплины (модуля)
15. Приложение №3. Перечень информационных технологий, используемых для освоения дисциплины (модуля), включая перечень программного обеспечения и информационных справочных систем

Программу дисциплины разработал(а)(и) доцент, к.н. Насрутдинов М.Ф. (кафедра компьютерной математики и информатики, отделение педагогического образования), Marat.Nasrutdinov@kpfu.ru

1. Перечень планируемых результатов обучения по дисциплине (модулю), соотнесенных с планируемыми результатами освоения ОПОП ВО

Обучающийся, освоивший дисциплину (модуль), должен обладать следующими компетенциями:

Шифр компетенции	Расшифровка приобретаемой компетенции
ОПК-1	Способен находить, формулировать и решать актуальные и значимые проблемы прикладной и компьютерной математики
ОПК-2	Способен создавать и исследовать новые математические модели в естественных науках, совершенствовать и разрабатывать концепции, теории и методы
ОПК-3	Способен самостоятельно создавать прикладные программные средства на основе современных информационных технологий и сетевых ресурсов, в том числе отечественного производства

Обучающийся, освоивший дисциплину (модуль):

Должен демонстрировать способность и готовность:

- понимать принципы, методы и подходы машинного обучения;
- применять теоретические знания, и практические умения и навыки в области исследования задач анализа данных и их решения методами машинного обучения.

2. Место дисциплины (модуля) в структуре ОПОП ВО

Данная дисциплина (модуль) включена в раздел "Б1.О.08 Дисциплины (модули)" основной профессиональной образовательной программы 02.04.01 "Математика и компьютерные науки (Методы математического и алгоритмического моделирования общенаучных и прикладных задач)" и относится к обязательным дисциплинам. Осваивается на 1 курсе в 1 семестре.

3. Объем дисциплины (модуля) в зачетных единицах с указанием количества часов, выделенных на контактную работу обучающихся с преподавателем (по видам учебных занятий) и на самостоятельную работу обучающихся

Общая трудоемкость дисциплины составляет 5 зачетных(ые) единиц(ы) на 180 часа(ов).

Контактная работа - 48 часа(ов), в том числе лекции - 24 часа(ов), практические занятия - 24 часа(ов), лабораторные работы - 0 часа(ов), контроль самостоятельной работы - 0 часа(ов).

Самостоятельная работа - 96 часа(ов).

Контроль (зачёт / экзамен) - 36 часа(ов).

Форма промежуточного контроля дисциплины: экзамен в 1 семестре.

4. Содержание дисциплины (модуля), структурированное по темам (разделам) с указанием отведенного на них количества академических часов и видов учебных занятий

4.1 Структура и тематический план контактной и самостоятельной работы по дисциплине (модулю)

N	Разделы дисциплины / модуля	Семестр	Виды и часы контактной работы, их трудоемкость (в часах)			Самостоятельная работа
			Лекции	Практические занятия	Лабораторные работы	
1.	Тема 1. Введение	1	4	4	0	12
2.	Тема 2. Задача классификации данных (логические и метрические)	1	4	4	0	21
3.	Тема 3. Задача классификации данных (линейные методы и Байесовская классификация)	1	4	4	0	21

N	Разделы дисциплины / модуля	Семестр	Виды и часы контактной работы, их трудоемкость (в часах)			Самостоятельная работа
			Лекции	Практические занятия	Лабораторные работы	
4.	Тема 4. Задача регрессии	1	6	6	0	21
5.	Тема 5. Нейронные сети	1	6	6	0	21
	Итого		24	24	0	96

4.2 Содержание дисциплины (модуля)

Тема 1. Введение

Введение. Анализ данных и машинное обучение. Исторический обзор теории машинного обучения. Обзор задач машинного обучения. Понятия и терминология машинного обучения. Классификация задач машинного обучения. Обучение с учителем. Метод коррекции ошибки. Метод обратного распространения ошибки. Обучение без учителя. Обучение с подкреплением. Метод ближайших соседей. Библиотеки языка Python для работы с данными (NumPy, Pandas, Scikit-Learn).

Тема 2. Задача классификации данных (логические и метрические)

Задачи классификации. Логические методы классификации. Решающие деревья. Жадная нисходящая стратегия 'разделяй и властвуй'. Алгоритм ID3. Недостатки жадной стратегии и способы их устранения. Проблема переобучения. Случайный лес. Метрические методы классификации. Гипотезы компактности и непрерывности. Обобщенный метрический классификатор. Метод ближайших соседей kNN и его обобщения. Подбор числа k по критерию скользящего контроля. Метод потенциальных функций и его связь с линейной моделью классификации. Непараметрическая регрессия. Локально взвешенный метод наименьших квадратов. Ядерное сглаживание. Задача отбора признаков. Жадный алгоритм построения метрики.

Тема 3. Задача классификации данных (линейные методы и Байесовская классификация)

Линейные методы классификации. Метод стохастического градиента для настройки линейных классификаторов. Регуляризация. Линейная регрессия от одной переменной. Модель представления. Функция оценивания. Метод градиентного спуска. Линейная регрессия от многих переменных. Полиномиальная регрессия. Проблема переобучения и недообучения. Отображение данных. Байесовская классификация и оценивание плотности

Тема 4. Задача регрессии

Задачи регрессии. Метод опорных векторов и логистическая регрессия. Оптимальная разделяющая гиперплоскость. Понятие зазора между классами. Случаи линейной делимости и отсутствия линейной делимости. Связь с минимизацией регуляризованного эмпирического риска. Кусочно-линейная функция потерь. Задача квадратичного программирования и двойственная задача. Понятие опорных векторов. Линейная регрессия. Задача регрессии, многомерная линейная регрессия. Метод наименьших квадратов, его вероятностный смысл и геометрический смысл. Понижение размерности и метод главных компонент. Композиции алгоритмов.

Тема 5. Нейронные сети

Нейронные сети. Градиентные методы оптимизации. Биологический нейрон, модель МакКаллока-Питтса как линейный классификатор. Функции активации. Проблема полноты. Задача исключаящего или. Полнота двухслойных сетей в пространстве булевых функций. Алгоритм обратного распространения ошибок. Эвристики для формирования начального приближения. Метод послонной настройки сети. Нейронные сети глубокого обучения (обзор).

5. Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине (модулю)

Самостоятельная работа обучающихся выполняется по заданию и при методическом руководстве преподавателя, но без его непосредственного участия. Самостоятельная работа подразделяется на самостоятельную работу на аудиторных занятиях и на внеаудиторную самостоятельную работу. Самостоятельная работа обучающихся включает как полностью самостоятельное освоение отдельных тем (разделов) дисциплины, так и проработку тем (разделов), осваиваемых во время аудиторной работы. Во время самостоятельной работы обучающиеся читают и конспектируют учебную, научную и справочную литературу, выполняют задания, направленные на закрепление знаний и отработку умений и навыков, готовятся к текущему и промежуточному контролю по дисциплине.

Организация самостоятельной работы обучающихся регламентируется нормативными документами, учебно-методической литературой и электронными образовательными ресурсами, включая:

Порядок организации и осуществления образовательной деятельности по образовательным программам высшего образования - программам бакалавриата, программам специалитета, программам магистратуры (утвержден приказом Министерства образования и науки Российской Федерации от 5 апреля 2017 года №301)

Письмо Министерства образования Российской Федерации №14-55-996ин/15 от 27 ноября 2002 г. "Об активизации самостоятельной работы студентов высших учебных заведений"

Устав федерального государственного автономного образовательного учреждения "Казанский (Приволжский) федеральный университет"

Правила внутреннего распорядка федерального государственного автономного образовательного учреждения высшего профессионального образования "Казанский (Приволжский) федеральный университет"

Локальные нормативные акты Казанского (Приволжского) федерального университета

6. Фонд оценочных средств по дисциплине (модулю)

Фонд оценочных средств по дисциплине (модулю) включает оценочные материалы, направленные на проверку освоения компетенций, в том числе знаний, умений и навыков. Фонд оценочных средств включает оценочные средства текущего контроля и оценочные средства промежуточной аттестации.

В фонде оценочных средств содержится следующая информация:

- соответствие компетенций планируемым результатам обучения по дисциплине (модулю);
- критерии оценивания сформированности компетенций;
- механизм формирования оценки по дисциплине (модулю);
- описание порядка применения и процедуры оценивания для каждого оценочного средства;
- критерии оценивания для каждого оценочного средства;
- содержание оценочных средств, включая требования, предъявляемые к действиям обучающихся, демонстрируемым результатам, задания различных типов.

Фонд оценочных средств по дисциплине находится в Приложении 1 к программе дисциплины (модулю).

7. Перечень литературы, необходимой для освоения дисциплины (модуля)

Освоение дисциплины (модуля) предполагает изучение основной и дополнительной учебной литературы.

Литература может быть доступна обучающимся в одном из двух вариантов (либо в обоих из них):

- в электронном виде - через электронные библиотечные системы на основании заключенных КФУ договоров с правообладателями;

- в печатном виде - в Научной библиотеке им. Н.И. Лобачевского. Обучающиеся получают учебную литературу на абонементе по читательским билетам в соответствии с правилами пользования Научной библиотекой.

Электронные издания доступны дистанционно из любой точки при введении обучающимся своего логина и пароля от личного кабинета в системе "Электронный университет". При использовании печатных изданий библиотечный фонд должен быть укомплектован ими из расчета не менее 0,5 экземпляра (для обучающихся по ФГОС 3++ - не менее 0,25 экземпляра) каждого из изданий основной литературы и не менее 0,25 экземпляра дополнительной литературы на каждого обучающегося из числа лиц, одновременно осваивающих данную дисциплину.

Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины (модуля), находится в Приложении 2 к рабочей программе дисциплины. Он подлежит обновлению при изменении условий договоров КФУ с правообладателями электронных изданий и при изменении комплектования фондов Научной библиотеки КФУ.

8. Перечень ресурсов информационно-телекоммуникационной сети "Интернет", необходимых для освоения дисциплины (модуля)

Наборы данных для машинного обучения - <http://archive.ics.uci.edu/ml/datasets.html>

Онлайн курс Andrew Ng - www.coursera.org/learn/machine-learning

Ресурс, посвященный машинному обучению - <http://www.machinelearning.ru/>

9. Методические указания для обучающихся по освоению дисциплины (модуля)

Вид работ	Методические рекомендации
лекции	Студентам необходимо посещать лекции и вести конспект лекций вслед за изложением материала преподавателем. Рекомендуется прорабатывать конспект в течение дня после лекции и просматривать его вновь накануне следующей лекции. В случае обнаружения ошибок или возникновения вопросов по предыдущему материалу необходимо обратиться к преподавателю.

Вид работ	Методические рекомендации
практические занятия	Для подготовки к практическим занятиям студенту рекомендуется предварительно проработать как лекционный материал, так и материал предыдущих практических занятий. Основой для подготовки служит добросовестное выполнение домашнего задания. Для успешного решения задач первой части курса студентам рекомендуется вспомнить материал, освоенный в предыдущих семестрах в рамках базовых математических дисциплин.
самостоятельная работа	Самостоятельная работа студентов состоит из двух основных частей - проработка лекционного материала и выполнения домашних заданий. Для освоения теоретического и практического материала, в случае, когда конспектов оказывается недостаточным, или для более детальной проработки отдельных тем рекомендуется использовать литературу, указанную в соответствующем разделе. Все возникающие вопросы рекомендуется заранее четко сформулировать и впоследствии обсудить с преподавателем.
экзамен	Залогом успешной сдачи экзамена является работа в течение всего семестра. Непосредственную подготовку к экзамену рекомендуется разделить на два этапа. На первом этапе прорабатываются все экзаменационные вопросы и формулируются вопросы к преподавателю в рамках консультации по разделам, недостаточно подробно описанным в рамках лекционного курса или более трудным в освоении материала. После консультации происходит окончательная проработка и закрепление материала по всем экзаменационным вопросам.

10. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине (модулю), включая перечень программного обеспечения и информационных справочных систем (при необходимости)

Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине (модулю), включая перечень программного обеспечения и информационных справочных систем, представлен в Приложении 3 к рабочей программе дисциплины (модуля).

11. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине (модулю)

Материально-техническое обеспечение образовательного процесса по дисциплине (модулю) включает в себя следующие компоненты:

Помещения для самостоятельной работы обучающихся, укомплектованные специализированной мебелью (столы и стулья) и оснащенные компьютерной техникой с возможностью подключения к сети "Интернет" и обеспечением доступа в электронную информационно-образовательную среду КФУ.

Учебные аудитории для контактной работы с преподавателем, укомплектованные специализированной мебелью (столы и стулья).

Компьютер и принтер для распечатки раздаточных материалов.

Мультимедийная аудитория.

Компьютерный класс.

12. Средства адаптации преподавания дисциплины к потребностям обучающихся инвалидов и лиц с ограниченными возможностями здоровья

При необходимости в образовательном процессе применяются следующие методы и технологии, облегчающие восприятие информации обучающимися инвалидами и лицами с ограниченными возможностями здоровья:

- создание текстовой версии любого нетекстового контента для его возможного преобразования в альтернативные формы, удобные для различных пользователей;
- создание контента, который можно представить в различных видах без потери данных или структуры, предусмотреть возможность масштабирования текста и изображений без потери качества, предусмотреть доступность управления контентом с клавиатуры;
- создание возможностей для обучающихся воспринимать одну и ту же информацию из разных источников - например, так, чтобы лица с нарушениями слуха получали информацию визуально, с нарушениями зрения - аудиально;
- применение программных средств, обеспечивающих возможность освоения навыков и умений, формируемых дисциплиной, за счёт альтернативных способов, в том числе виртуальных лабораторий и симуляционных технологий;
- применение дистанционных образовательных технологий для передачи информации, организации различных форм интерактивной контактной работы обучающегося с преподавателем, в том числе вебинаров, которые могут быть использованы для проведения виртуальных лекций с возможностью взаимодействия всех участников дистанционного обучения, проведения семинаров, выступления с докладами и защиты выполненных работ, проведения тренингов, организации коллективной работы;

- применение дистанционных образовательных технологий для организации форм текущего и промежуточного контроля;
- увеличение продолжительности сдачи обучающимся инвалидом или лицом с ограниченными возможностями здоровья форм промежуточной аттестации по отношению к установленной продолжительности их сдачи:
- продолжительности сдачи зачёта или экзамена, проводимого в письменной форме, - не более чем на 90 минут;
- продолжительности подготовки обучающегося к ответу на зачёте или экзамене, проводимом в устной форме, - не более чем на 20 минут;
- продолжительности выступления обучающегося при защите курсовой работы - не более чем на 15 минут.

Программа составлена в соответствии с требованиями ФГОС ВО и учебным планом по направлению 02.04.01 "Математика и компьютерные науки" и магистерской программе "Методы математического и алгоритмического моделирования общенаучных и прикладных задач".

Перечень литературы, необходимой для освоения дисциплины (модуля)

Направление подготовки: 02.04.01 - Математика и компьютерные науки

Профиль подготовки: Методы математического и алгоритмического моделирования общенаучных и прикладных задач

Квалификация выпускника: магистр

Форма обучения: очное

Язык обучения: русский

Год начала обучения по образовательной программе: 2018

Основная литература:

1. Коэльо Луис Педро, Построение систем машинного обучения на языке Python [Электронный ресурс] / Луис Педро Коэльо, Вилли Ричарт - М. : ДМК Пресс, 2016. - 302 с. - ISBN 978-5-97060-330-7 - URL: <http://www.studentlibrary.ru/book/ISBN9785970603307.html>
2. Флах П., Машинное обучение. Наука и искусство построения алгоритмов, которые извлекают знания из данных [Электронный ресурс] / Флах П. - М. : ДМК Пресс, 2015. - 400 с. - ISBN 978-5-97060-273-7 - URL: <http://www.studentlibrary.ru/book/ISBN9785970602737.html>
3. Маккинли У., Python и анализ данных [Электронный ресурс] / Уэс Маккинли - М. : ДМК Пресс, 2015. - 482 с. - ISBN 978-5-97060-315-4 - URL: <http://www.studentlibrary.ru/book/ISBN9785970603154.html>
4. Вьюгин, В.В. Математические основы машинного обучения и прогнозирования : учебное пособие / В.В. Вьюгин. - Москва : МЦНМО, 2014. - 304 с. - ISBN 978-5-4439-2014-6. - Текст : электронный // Электронно-библиотечная система 'Лань' : [сайт]. - URL: <https://e.lanbook.com/book/56397>

Дополнительная литература:

1. Саммерфилд М., Python на практике [Электронный ресурс] / Марк Саммерфилд - М. : ДМК Пресс, 2014. - 338 с. - ISBN 978-5-97060-095-5 - URL: <http://www.studentlibrary.ru/book/ISBN9785970600955.html>
2. Осипов Г.С., Методы искусственного интеллекта [Электронный ресурс] / Осипов Г.С. - М. : ФИЗМАТЛИТ, 2011. - 296 с. - ISBN 978-5-9221-1323-6 - URL: <http://www.studentlibrary.ru/book/ISBN9785922113236.html>

Приложение 3
к рабочей программе дисциплины (модуля)
Б1.О.08 Машинное обучение

Перечень информационных технологий, используемых для освоения дисциплины (модуля), включая перечень программного обеспечения и информационных справочных систем

Направление подготовки: 02.04.01 - Математика и компьютерные науки

Профиль подготовки: Методы математического и алгоритмического моделирования общенаучных и прикладных задач

Квалификация выпускника: магистр

Форма обучения: очное

Язык обучения: русский

Год начала обучения по образовательной программе: 2018

Освоение дисциплины (модуля) предполагает использование следующего программного обеспечения и информационно-справочных систем:

Операционная система Microsoft Windows 7 Профессиональная или Windows XP (Volume License)

Пакет офисного программного обеспечения Microsoft Office 365 или Microsoft Office Professional plus 2010

Браузер Mozilla Firefox

Браузер Google Chrome

Adobe Reader XI или Adobe Acrobat Reader DC

Kaspersky Endpoint Security для Windows

Учебно-методическая литература для данной дисциплины имеется в наличии в электронно-библиотечной системе "ZNANIUM.COM", доступ к которой предоставлен обучающимся. 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, учебно-методические комплексы, 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 образования (ФГОС ВО) нового поколения.

Учебно-методическая литература для данной дисциплины имеется в наличии в электронно-библиотечной системе Издательства "Лань", доступ к которой предоставлен обучающимся. 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.

Учебно-методическая литература для данной дисциплины имеется в наличии в электронно-библиотечной системе "Консультант студента", доступ к которой предоставлен обучающимся. Многопрофильный образовательный ресурс "Консультант студента" является электронной библиотечной системой (ЭБС), предоставляющей доступ через сеть Интернет к учебной литературе и дополнительным материалам, приобретенным на основании прямых договоров с правообладателями. Полностью соответствует требованиям федеральных государственных образовательных стандартов высшего образования к комплектованию библиотек, в том числе электронных, в части формирования фондов основной и дополнительной литературы.