

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное автономное образовательное учреждение высшего образования
"Казанский (Приволжский) федеральный университет"
Высшая школа информационных технологий и интеллектуальных систем



подписано электронно-цифровой подписью

Программа дисциплины

Машинное обучение

Направление подготовки: 09.03.04 - Программная инженерия

Профиль подготовки: Технологии разработки информационных систем

Квалификация выпускника: бакалавр

Форма обучения: очное

Язык обучения: русский

Год начала обучения по образовательной программе: 2016

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) профессор, к.н. Магид Е.А. (кафедра интеллектуальной робототехники, Высшая школа информационных технологий и интеллектуальных систем), magid@it.kfu.ru Сагитов А.Г.

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
ПК-12	способность к формализации в своей предметной области с учетом ограничений используемых методов исследования

Обучающийся, освоивший дисциплину (модуль):

Должен знать:

основные постановки задач анализа данных и методы их решения.

Должен уметь:

применять на практике алгоритмы решения задач интеллектуального анализа данных.

Должен владеть:

навыками разработки решений задач анализа данных применяя современные методы и инструменты.

Должен демонстрировать способность и готовность:

назвать и использовать алгоритмы, которые могут быть использованы при решении задач анализа данных;

провести очистку и подготовку данных для последующего анализа;

провести анализ с применением современного инструментария.

2. Место дисциплины (модуля) в структуре ОПОП ВО

Данная дисциплина (модуль) включена в раздел "Б1.В.ОД.11 Дисциплины (модули)" основной профессиональной образовательной программы 09.03.04 "Программная инженерия (Технологии разработки информационных систем)" и относится к обязательным дисциплинам.

Осваивается на 4 курсе в 7 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 6 зачетных(ые) единиц(ы) на 216 часа(ов).

Контактная работа - 72 часа(ов), в том числе лекции - 36 часа(ов), практические занятия - 36 часа(ов), лабораторные работы - 0 часа(ов), контроль самостоятельной работы - 0 часа(ов).

Самостоятельная работа - 90 часа(ов).

Контроль (зачёт / экзамен) - 54 часа(ов).

Форма промежуточного контроля дисциплины: экзамен в 7 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 в машинное обучение.	7	6	5	0	14
2.	Тема 2. Задачи регрессии.	7	6	8	0	12
3.	Тема 3. Задачи классификации.	7	6	8	0	18
4.	Тема 4. Задачи кластеризации.	7	6	6	0	18

N	Разделы дисциплины / модуля	Семестр	Виды и часы контактной работы, их трудоемкость (в часах)			Самостоятельная работа
			Лекции	Практические занятия	Лабораторные работы	
5.	Тема 5. Задачи ассоциации и рекомендации.	7	6	6	0	16
6.	Тема 6. Понижение размерности и ансамбли моделей.	7	6	3	0	12
	Итого		36	36	0	90

4.2 Содержание дисциплины (модуля)

Тема 1. Введение в машинное обучение.

Понятия и терминология машинного обучения. Классификация задач машинного обучения. Обучение с учителем. Метод коррекции ошибки. Метод обратного распространения ошибки.

Обучение без учителя. Альфа-система подкрепления. Гамма-система подкрепления.

Метод ближайших соседей.

Тема 2. Задачи регрессии.

Постановка задачи регрессии. Методы решения задачи регрессии. Линейная регрессия. Нейронные сети. Деревья решений. Оценка качества. Определение степени детерминированности вариации критериальной (зависимой) переменной предикторами (независимыми переменными). Предсказание значения зависимой переменной с помощью независимой(-ых).

Определение вклада отдельных независимых переменных в вариацию зависимой.

Тема 3. Задачи классификации.

Постановка задачи классификации. Методы решения задачи классификации. Логистическая регрессия. Оценки качества классификации. Двухклассовая классификация. Многоклассовая классификация. Непересекающиеся классы. Пересекающиеся классы. Нечёткие классы.

Признаковое описание.

Тема 4. Задачи кластеризации.

Постановка задачи кластеризации. Иерархические и агломеративные алгоритмы. Самоорганизующаяся карта Кохонена. Признаковое описание объектов. Матрица расстояний между объектами. Понимание данных путём выявления кластерной структуры. Сжатие данных. Обнаружение новизны. Метрика Хэмминга. Евклидова метрика. Взвешенная евклидова метрика. Метрика Минковского.

Тема 5. Задачи ассоциации и рекомендации.

Задачи ассоциации и рекомендации. Алгоритм *a priori*. Ассоциативные правила. Коллаборативные фильтрации. Статистическое оценивание. Критерии согласия.

Критерии сдвига. Критерии нормальности. Критерии однородности. Критерии симметричности. Критерии тренда, стационарности и случайности. Критерии выбросов.

Критерии дисперсионного анализа. Критерии корреляционного анализа.

Тема 6. Понижение размерности и ансамбли моделей.

Задача понижения размерности. Метод главных компонент. *t-SNE*. Методы агрегации и усиления моделей.

Бустинг и беггинг. Модель панельных данных с фиксированными эффектами. Объединённая модель панельных данных. Модель панельных данных со случайными эффектами. Модель панельных данных с временными эффектами.

Ротационная панель.

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)

Intro to Machine Learning Course | Udacity - <https://www.udacity.com/course/intro-to-machine-learning--ud120>

Introduction to Machine Learning - Online Course - DataCamp - <https://www.datacamp.com/courses/introduction-to-machine-learning-with-r>

Machine Learning | Coursera - <https://www.coursera.org/learn/machine-learning>

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

Вид работ	Методические рекомендации
лекции	В ходе лекционных занятий вести конспектирование учебного материала, задавая преподавателю уточняющие вопросы для разрешения спорных ситуаций. Обращать внимание содержание тех или иных явлений и процессов, научные выводы и практические рекомендации. Рекомендуется оставить в рабочих конспектах поля, на которых делать дополняющие материал

пометки, подчеркивать важность тех или иных тезисов.

Вид работ	Методические рекомендации
практические занятия	Необходимо изучить литературу, рекомендуемую для выполнения практического занятия. Студент должен выявить необходимость дополнительных источников и материалов. Студентам должен овладеть навыками и умениями использования теоретического знания применительно к особенностям изучаемой дисциплины. Подготовка к практическим занятиям предполагает предварительную самостоятельную работу студентов на основе методических разработок. Студент овладевает первоначальными профессиональными умениями и навыками в процессе практических занятий. Знания обобщаются, систематизируются, углубляются и конкретизируются теоретические вопросы, вырабатывается способность и готовность использовать теоретические знания. Работа студента представлена в виде экспериментальной работы, решения ситуационных задач, обсуждения проблемных вопросов, работы с компьютером и т.п. Результатом самостоятельной работы обучающихся на практическом занятии является письменный ответ.
самостоятельная работа	Необходимо изучить основную литературу, ознакомиться с дополнительной литературой. Учесть рекомендации преподавателя и требования учебной программы при доработке конспекта лекций, делая в нем соответствующие записи. Выделить вопросы из списка литературы, рекомендованной преподавателем и предусмотренной учебной программой. Формулирование выводов осуществляется в процессе творческой дискуссии, протекающей с соблюдением методологических требований к научному познанию.
экзамен	К экзамену допускается студент, выполнивший в полном объеме задания, указанные в рабочей программе. Студент самостоятельно выполняет и сдает на проверку в письменном виде общие или индивидуальные задания, определяемые преподавателем. Экзамен проходит в устной и письменной форме на основе перечня вопросов, указанных в действующей рабочей программы учебной дисциплины. На оценку влияют следующие критерии: степень раскрытия темы, стиль изложения и творческий подход к ответу на вопрос, правильность аргументация выводов и др. Студентам рекомендуется готовиться к экзамену в группе по три-четыре человека. Следует внимательно прочитать вопросы и составить план ответа на каждый вопрос, выделив ключевые моменты материала. Результаты сдачи экзаменов оцениваются отметкой "отлично", "хорошо", "удовлетворительно" или "неудовлетворительно".

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 09.03.04 "Программная инженерия" и профилю подготовки "Технологии разработки информационных систем".

Перечень литературы, необходимой для освоения дисциплины (модуля)

Направление подготовки: 09.03.04 - Программная инженерия

Профиль подготовки: Технологии разработки информационных систем

Квалификация выпускника: бакалавр

Форма обучения: очное

Язык обучения: русский

Год начала обучения по образовательной программе: 2016

Основная литература:

1. 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 - Режим доступа: <http://www.studentlibrary.ru/book/ISBN9785970602737.html>
2. Методы и алгоритмы обработки данных: учеб. пособие / А.А. Григорьев. - М.: ИНФРА-М, 2018. - 256 с. - (Высшее образование). - Режим доступа: <http://znanium.com/catalog/product/922736>
3. 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 - Режим доступа: <http://www.studentlibrary.ru/book/ISBN9785970603307.html>

Дополнительная литература:

1. Болотова Л.С., Системы искусственного интеллекта: модели и технологии, основанные на знаниях [Электронный ресурс]: учебник / Л.С. Болотова. - М.: Финансы и статистика, 2012. - 664 с. - ISBN 978-5-279-03530-4 - Режим доступа: <http://www.studentlibrary.ru/book/ISBN9785279035304.html>
2. Маккинли У., Python и анализ данных [Электронный ресурс] / Уэс Маккинли - М.: ДМК Пресс, 2015. - 482 с. - ISBN 978-5-97060-315-4 - Режим доступа: <http://www.studentlibrary.ru/book/ISBN9785970603154.html>
3. Боровиков В.П., Популярное введение в современный анализ данных в системе STATISTICA [Электронный ресурс]: Учебное пособие для вузов / Боровиков В.П. - М.: Горячая линия - Телеком, 2013. - 288 с. - ISBN 978-5-9912-0326-5 - Режим доступа: <http://www.studentlibrary.ru/book/ISBN9785991203265.html>

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

Направление подготовки: 09.03.04 - Программная инженерия

Профиль подготовки: Технологии разработки информационных систем

Квалификация выпускника: бакалавр

Форма обучения: очное

Язык обучения: русский

Год начала обучения по образовательной программе: 2016

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 "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.