

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное автономное образовательное учреждение высшего образования
"Казанский (Приволжский) федеральный университет"
Институт геологии и нефтегазовых технологий



УТВЕРЖДАЮ

Проректор по образовательной деятельности КФУ

Проф. Д. А. Таюрский

» _____ 20__ г.

подписано электронно-цифровой подписью

Программа дисциплины

Машинные методы обучения

Направление подготовки: 05.04.01 - Геология

Профиль подготовки: Перспективные геоинформационные технологии в геологии и геофизике

Квалификация выпускника: магистр

Форма обучения: очное

Язык обучения: русский

Год начала обучения по образовательной программе: 2017

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. (доцент) Чернова И.Ю. (кафедра геофизики и геоинформационных технологий, Институт геологии и нефтегазовых технологий), Inna.Chernova@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	способностью самостоятельно приобретать, осмысливать, структурировать и использовать в профессиональной деятельности новые знания и умения, развивать свои инновационные способности
ОПК-2	способностью самостоятельно формулировать цели исследований, устанавливать последовательность решения профессиональных задач
ОПК-4	способностью профессионально выбирать и творчески использовать современное научное и техническое оборудование для решения научных и практических задач
ОПК-5	способностью критически анализировать, представлять, защищать, обсуждать и распространять результаты своей профессиональной деятельности
ПК-1	способностью формировать диагностические решения профессиональных задач путем интеграции фундаментальных разделов геологических наук и специализированных знаний, полученных при освоении программы магистратуры
ПК-8	готовностью к проектированию комплексных научно-исследовательских и научно-производственных работ при решении профессиональных задач

Обучающийся, освоивший дисциплину (модуль):

Должен знать:

основные принципы создания и область применения машинных методов обучения

Должен уметь:

ориентироваться в применении машинных методов обучения

Должен владеть:

теоретическими знаниями об используемом машинных методах обучения

Должен демонстрировать способность и готовность:

применять полученные навыки и знания в практической деятельности

2. Место дисциплины (модуля) в структуре ОПОП ВО

Данная дисциплина (модуль) включена в раздел "Б1.В.ДВ.4 Дисциплины (модули)" основной профессиональной образовательной программы 05.04.01 "Геология (Перспективные геоинформационные технологии в геологии и геофизике)" и относится к дисциплинам по выбору.

Осваивается на 1 курсе в 2 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 2 зачетных(ые) единиц(ы) на 72 часа(ов).

Контактная работа - 28 часа(ов), в том числе лекции - 8 часа(ов), практические занятия - 0 часа(ов), лабораторные работы - 20 часа(ов), контроль самостоятельной работы - 0 часа(ов).

Самостоятельная работа - 44 часа(ов).

Контроль (зачёт / экзамен) - 0 часа(ов).

Форма промежуточного контроля дисциплины: зачет во 2 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. Задачи обучения по прецедентам (примерам).	2	1	0	4	8
2.	Тема 2. Байесовские методы обучения.	2	1	0	4	8
3.	Тема 3. Метрические методы обучения.	2	2	0	4	8
4.	Тема 4. Линейные методы обучения.	2	2	0	4	8
5.	Тема 5. Методы восстановления зависимостей.	2	2	0	4	12
	Итого		8	0	20	44

4.2 Содержание дисциплины (модуля)

Тема 1. Задачи обучения по прецедентам (примерам).

Основные понятия и определения. Объекты и признаки, степень измеримости, кодирование (контрасты). общие зависимости, закономерности, взаимосвязи.

Типология задач обучения по прецедентам. Обучение с учителем (supervised learning). Обучение без учителя (unsupervised learning). Частичное обучение (semi-supervised learning). Трансдуктивное обучение (transductive learning). Обучение с подкреплением (reinforcement learning).

Тема 2. Байесовские методы обучения.

Вероятностная постановка задачи классификации: функционал среднего риска, оптимальное байесовское решающее правило.

Лекция по байесовским методам решения различных задач машинного обучения (классификации, восстановления регрессии, уменьшения размерности, разделения смесей, тематического моделирования и др.), которые в настоящее время активно развиваются в мире. Большинство современных научных публикаций по машинному обучению используют вероятностное моделирование, опирающееся на байесовский подход к теории вероятностей. Последний позволяет эффективно учитывать различные предпочтения пользователя при построении решающих правил прогноза. Кроме того, он позволяет решать задачи выбора структурных параметров модели. В частности, здесь удается решать без комбинаторного перебора задачи селекции признаков, выбора числа кластеров в данных, размерности редуцированного пространства при уменьшении размерности, значений коэффициентов регуляризации и пр. В байесовском подходе вероятность интерпретируется как мера незнания, а не как объективная случайность. Простые правила оперирования с вероятностью, такие как формула полной вероятности и формула Байеса, позволяют проводить рассуждения в условиях неопределенности. В этом смысле байесовский подход к теории вероятностей можно рассматривать как обобщение классической булевой логики.

Тема 3. Метрические методы обучения.

Метод ближайшего соседа и его обобщения: обобщенный метрический классификатор, метод ближайших соседей

Задачи классификации и регрессии. Гипотеза компактности (для классификации). Гипотеза непрерывности (для регрессии). Формализация понятия "близости". Евклидово расстояние и его обобщение. Обобщенный алгоритм. Lazy learning. Метод k ближайших соседей. Оптимизация k

Тема 4. Линейные методы обучения.

Градиентные методы обучения

Минимизация эмпирического риска

Линейный классификатор

Метод стохастического градиента

Регуляризация эмпирического риска

2 Логистическая регрессия (LR)

Экспонентные семейства плотностей

Основная теорема

Принцип максимума правдоподобия

Скоринг

3 Балансировка ошибок и ROC-кривая

Постановка задачи

Определение ROC-кривой

Эффективное построение ROC-кривой

Градиентная максимизация AUC

4 Метод опорных векторов (SVM)

Принцип оптимальной разделяющей гиперплоскости

Двойственная задача

Ядра и спрямляющие пространства

Метод релевантных векторов (RVM)

Тема 5. Методы восстановления зависимостей.

Метод наименьших квадратов. Непараметрическая регрессия: ядерное сглаживание: формула Надарая-Ватсона, выбор ядра и ширины окна, проблема выбросов: робастная непараметрическая регрессия

Построение и оценивание иерархических вероятностных моделей в типовых задачах восстановления зависимостей по эмпирическим данным. Общие принципы построения иерархических вероятностных моделей в задачах восстановления зависимостей. Построение иерархических вероятностных моделей для случая представления независимых и целевой переменной в различных шкалах. Вероятностная модель для задачи линейной регрессии при представлении объектов потенциальными функциями для задачи оценивания положения сайта в результатах поискового запроса. Квази-вероятностная модель для задачи распознавания образов на два класса. Квази-вероятностная модель для оценивания порядковой регрессии на примере задачи оценивания положения сайта в результатах поискового запроса. Квази-вероятностная модель для оценивания регрессионной модели Кокса на примере задачи анализа продолжительности жизни.

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)

Департамент машинного обучения CMU - <http://www.ml.cmu.edu/>

Журнал "Машинное обучение?", MIT - <http://jmlr.csail.mit.edu/>

Курс машинного обучения в Стэнфордском университете - <http://cs229.stanford.edu/>

Профессиональный информационно-аналитический ресурс, посвященный машинному обучению, распознаванию образов и интеллектуальному анализу данных. - <http://www.machinelearning.ru/>

Статистическая система R - <http://cran.r-project.org/>

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

Вид работ	Методические рекомендации
лекции	В ходе лекционных занятий вести конспектирование учебного материала. Обращать внимание на категории, формулировки, раскрывающие содержание тех или иных явлений и процессов, научные выводы и практические рекомендации. Желательно оставить в рабочих конспектах поля, на которых делать пометки из рекомендованной литературы, дополняющие материал прослушанной лекции, а также подчеркивающие особую важность тех или иных теоретических положений. Задавать преподавателю уточняющие вопросы с целью уяснения теоретических положений, разрешения спорных ситуаций.
лабораторные работы	При выполнении практических заданий студент руководствуется правилами, изложенными преподавателем при постановке задачи на занятии и в описании работы. Кроме того, должен активно использоваться материал, изложенный на лекциях, и привлекаться дополнительная специальная литература. Студент самостоятельно анализирует полученные результаты, т.е. выполняет элементы научного поиска, на основе которого составляется письменный отчет. Этот отчет по своей форме должен содержать следующие разделы: краткую теоретическую часть, расчётный раздел, подробный анализ результатов, выводы, т.е. соответствовать структуре научно-технического отчета, научной статьи. Все необходимые (указанные в задании) графики должны быть выполнены в виде компьютерных рисунков с помощью программы компьютерной графики. Если при проверке отчёта преподавателем будут выявлены отклонения от установленных требований или ошибки, он должен быть доработан.
самостоятельная работа	Самостоятельная внеаудиторная работа является высшей формой самоорганизации познавательной деятельности студента и решает разнообразные дидактические задачи: закрепление, углубление, расширение, систематизация знаний, полученных во время внеаудиторных занятий, самостоятельное овладение новым учебным материалом, формирование умений и навыков самостоятельного умственного труда, профессиональных умений; развитие самостоятельности мышления, формирование волевых черт характера, способности к самоорганизации. Самостоятельная работа студентов требует определенного уровня способности к самообразованию, а также устойчивых навыков работы с учебной и научной литературой.

Вид работ	Методические рекомендации
зачет	<p>Зачет как форма контроля и организации обучения служит приемом проверки степени усвоения учебного материала и лекционных занятий, сформированных умений и навыков. По решению преподавателя зачет может быть выставлен без опроса ? по результатам работы обучающегося на лекционных и(или) практических занятиях.</p> <p>В период подготовки к зачету обучающиеся вновь обращаются к пройденному учебному материалу. Подготовка обучающегося к зачету включает в себя три этапа:</p> <ul style="list-style-type: none"> * самостоятельная работа в течение процесса обучения; * непосредственная подготовка в дни, предшествующие зачету по темам курса; * подготовка к ответу на вопросы, содержащиеся в программе дисциплин. <p>Литература для подготовки к зачету рекомендуется преподавателем.</p> <p>По окончании ответа преподаватель может задать обучающемуся дополнительные и уточняющие вопросы. На подготовку к ответу обучающемуся дается 30 минут с момента получения им вопросов.</p>

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.

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 05.04.01 "Геология" и магистерской программе "Перспективные геоинформационные технологии в геологии и геофизике".

Перечень литературы, необходимой для освоения дисциплины (модуля)

Направление подготовки: 05.04.01 - Геология

Профиль подготовки: Перспективные геоинформационные технологии в геологии и геофизике

Квалификация выпускника: магистр

Форма обучения: очное

Язык обучения: русский

Год начала обучения по образовательной программе: 2017

Основная литература:

Экспертные системы САПР: Учебное пособие / Ездаков А.Л. - М.:ИД ФОРУМ, НИЦ ИНФРА-М, 2016. - 160 с.: 60х90 1/16. - (Высшее образование) (Переплёт 7БЦ) ISBN 978-5-8199-0398-8
<http://znanium.com/bookread2.php?book=518395>

Информационные технологии и системы: Учебное пособие / Е.Л. Федотова. - М.: ИД ФОРУМ: НИЦ Инфра-М, 2013. - 352 с.: ил.; 60х90 1/16. - (Высшее образование). (переплет) ISBN 978-5-8199-0376-6
<http://znanium.com/bookread2.php?book=374014>

Компьютерные науки. Деревья, операционные системы, сети / И.Ф. Астахова, И.К. Астанин, И.Б. Крыжко. - М.: ФИЗМАТЛИТ, 2013. - 88 с.: 60х90 1/16. (обложка) ISBN 978-5-9221-1449-3, 500 экз.
<http://znanium.com/bookread2.php?book=428176>

Дополнительная литература:

Введение в геоинформационные системы: Учебное пособие / Я.Ю. Блиновская, Д.С. Задоя. - М.: Форум: НИЦ Инфра-М, 2013. - 112 с.: 60х90 1/16. - (Высшее образование: Бакалавриат). (обложка) ISBN 978-5-91134-698-0, 300 экз. <http://znanium.com/bookread2.php?book=372170>

Информационные системы: Учебное пособие / О.Л. Голицына, Н.В. Максимов, И.И. Попов. - 2-е изд. - М.: Форум: НИЦ ИНФРА-М, 2014. - 448 с.: ил.; 60х90 1/16. - (Высшее образование). (переплет) ISBN 978-5-91134-833-5, 1000 экз. <http://znanium.com/bookread2.php?book=435900>

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. - (Университетская серия). - ISBN 978-5-4257-0092-6. <http://znanium.com/bookread2.php?book=451186>

Приложение 3
к рабочей программе дисциплины (модуля)
Б1.В.ДВ.4 Машинные методы обучения

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

Направление подготовки: 05.04.01 - Геология

Профиль подготовки: Перспективные геоинформационные технологии в геологии и геофизике

Квалификация выпускника: магистр

Форма обучения: очное

Язык обучения: русский

Год начала обучения по образовательной программе: 2017

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