

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



УТВЕРЖДАЮ

Проректор по образовательной деятельности КФУ

Проф. Д. А. Таюрский

» _____ 20__ г.

подписано электронно-цифровой подписью

Программа дисциплины

Дополнительные главы машинного обучения

Направление подготовки: 01.04.02 - Прикладная математика и информатика

Профиль подготовки: Анализ данных и его приложения

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) заведующий кафедрой, д.н. (доцент) Миссаров М.Д. (кафедра анализа данных и исследования операций, отделение фундаментальной информатики и информационных технологий), Moukadas.Missarov@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
ОК-1	способность к абстрактному мышлению, анализу, синтезу
ОК-3	готовность к саморазвитию, самореализации, использованию творческого потенциала
ОПК-3	способностью самостоятельно приобретать с помощью информационных технологий и использовать в практической деятельности новые знания и умения, в том числе, в новых областях знаний, непосредственно не связанных со сферой деятельности, расширять и углублять своё научное мировоззрение
ОПК-4	способностью использовать и применять углубленные знания в области прикладной математики и информатики
ПК-8	способностью разрабатывать корпоративные стандарты и профили функциональной стандартизации приложений, систем, информационной инфраструктуры
ПК-9	способностью к преподаванию математических дисциплин и информатики в образовательных организациях основного общего, среднего общего, среднего профессионального и высшего образования

Обучающийся, освоивший дисциплину (модуль):

Должен знать:

- основные постановки современных задач машинного обучения;
- основы анализа главных компонент и дискриминантного анализа ;
- методы оценивания и комбинирования классификаторов, логистической регрессии, нейронных сетей с различной архитектурой.

Должен уметь:

- формализовать задачи в различных прикладных областях на основе математических моделей дискриминантного анализа и машины опорных векторов, методов главных компонент и логистической регрессии;

- применять методы и алгоритмы машинного обучения при решении прикладных задач анализа данных;

Должен владеть:

- современным математическим аппаратом анализа данных;
- навыками постановки научно-исследовательских задач.

Должен демонстрировать способность и готовность:

- осуществлять поиск информации и предварительную обработку данных по полученному заданию, подбор адекватных методов машинного обучения, необходимых для решения поставленных задач
- применять в профессиональной деятельности знания, умения, навыки, полученные в ходе освоения дисциплины

2. Место дисциплины (модуля) в структуре ОПОП ВО

Данная дисциплина (модуль) включена в раздел "Б1.В.ОД.5 Дисциплины (модули)" основной профессиональной образовательной программы 01.04.02 "Прикладная математика и информатика (Анализ данных и его приложения)" и относится к обязательным дисциплинам.

Осваивается на 2 курсе в 3 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 4 зачетных(ые) единиц(ы) на 144 часа(ов).

Контактная работа - 42 часа(ов), в том числе лекции - 24 часа(ов), практические занятия - 18 часа(ов), лабораторные работы - 0 часа(ов), контроль самостоятельной работы - 0 часа(ов).

Самостоятельная работа - 66 часа(ов).

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

Форма промежуточного контроля дисциплины: экзамен в 3 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. Снижение размерности	3	3	2	0	11
2.	Тема 2. Методы классификации	3	7	4	0	16
3.	Тема 3. Машины опорных векторов	3	6	2	0	17
4.	Тема 4. Оценивание и комбинирование классификаторов	3	6	2	0	11
5.	Тема 5. Глубокое обучение.	3	2	8	0	11
	Итого		24	18	0	66

4.2 Содержание дисциплины (модуля)

Тема 1. Снижение размерности

Матрица данных и матрица ковариаций. Собственные числа и собственные векторы положительно-определенной матрицы. Построение первой главной компоненты. Свойство проекции на первую главную компоненту. Построение второй главной компоненты. Алгоритм понижения размерности методом главных компонент (общий случай). Выбор подмножеств.

Тема 2. Методы классификации

Многомерное гауссовское распределение. Статистическое оценивание параметров многомерного гауссовского распределения. Минимизация математического ожидания цены ошибочной классификации. Полная вероятность ошибочной классификации. Связь алгоритма минимизации полной вероятности ошибочной классификации с алгоритмом байесовской классификации. Классификация двух нормальных выборок с одинаковыми матрицами ковариаций. Классификация двух нормальных выборок с разными матрицами ковариаций. Линейный дискриминант Фишера. Логистическая модель. Оценивание коэффициентов логистической регрессии. Пробит-модель.

Тема 3. Машины опорных векторов

Метод опорных векторов с максимальным зазором. Метод опорных векторов с мягким зазором. Теорема о представителе. Ядерный трюк. Машина опорных векторов.

Эффективность ядерного трюка. Теорема Мерсера. Построение ядерных функций. Ядерные функции для строк и графов. Машины опорных векторов для случаев с несколькими классами. Регрессия и метод опорных векторов.

Тема 4. Оценивание и комбинирование классификаторов

Оценивание качества классификации. Матрица ошибок бинарной классификации. ROC-анализ. Перекрестная проверка. Бутстрап. Минимизация средних потерь. Разложение смещение-дисперсия. Комбинирование классификаторов (математические мотивировки). Процедуры комбинирования классификаторов. Баггинг. Бустинг. Случайные леса.

Тема 5. Глубокое обучение.

Глубокое обучение нейронных сетей. Стохастический градиентный спуск. Способы регуляризации.

Проблемы оптимизации нейронных сетей. Сверточные нейронные сети. Операция свертки. Пулинг. Рекуррентные нейронные сети. Автокодировщики. Примеры применения глубоких нейронных сетей в задачах обработки текстов и изображений.

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

Вьюгин, В.В. Математические основы машинного обучения и прогнозирования [Электронный ресурс] : учеб. пособие - Москва : МЦНМО, 2013. 304 с. - <https://e.lanbook.com/book/56397>

Машинное обучение (курс лекций С.К.В.Воронцов) -

[http://www.machinelearning.ru/wiki/index.php?title=Машинное_обучение_\(курс_лекций%2C_К.В.Воронцов\)](http://www.machinelearning.ru/wiki/index.php?title=Машинное_обучение_(курс_лекций%2C_К.В.Воронцов))

Флах, П. Машинное обучение. Наука и искусство построения алгоритмов, которые извлекают знания из данных [Электронный ресурс] , Москва : ДМК Пресс, 2015. 400 с. - <https://e.lanbook.com/book/69955>

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)

Байесовские_методы_машинного_обучения_(курс_лекций)/_2017 Д.П. Ветров - <http://www.machinelearning.ru/wiki/index.php?title=>

Байесовские_методы_машинного_обучения_(курс_лекций)/_2017 Д.П. Ветров

Машинное обучение (курс лекций, Н.Ю.Золотых) - <http://www.uic.unn.ru/~zny/ml/>

Машинное_обучение_(курс_лекций_С_К.В.Воронцов) -

[http://www.machinelearning.ru/wiki/index.php?title=Машинное_обучение_\(курс_лекций%2C_К.В.Воронцов\)](http://www.machinelearning.ru/wiki/index.php?title=Машинное_обучение_(курс_лекций%2C_К.В.Воронцов))

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

Вид работ	Методические рекомендации
лекции	После прослушивания каждой лекции студент должен дома самостоятельно ее проработать, используя конспект и ту основную и дополнительную литературу, которая была рекомендована преподавателем. Такая самостоятельная работа даст возможность студенту на очередной лекции лучше понимать и усваивать новый материал.
практические занятия	Практическое занятие ? форма систематических учебных занятий, с помощью которых обучающиеся изучают тот или иной раздел определенной научной дисциплины, входящей в состав учебного плана. Для того чтобы практические занятия приносили максимальную пользу, необходимо помнить, что упражнение и решение задач проводятся по вычитанному на лекциях материалу и связаны, как правило, с детальным разбором отдельных вопросов лекционного курса. При самостоятельном решении задач нужно обосновывать каждый этап решения, исходя из теоретических положений курса. При подготовке к практическим занятиям следует использовать основную литературу из представленного списка, а также руководствоваться приведенными указаниями и рекомендациями. Для наиболее глубокого освоения дисциплины рекомендуется изучать литературу, обозначенную как ?дополнительная? в представленном списке. На практических занятиях приветствуется активное участие в обсуждении конкретных ситуаций, способность на основе полученных знаний находить наиболее эффективные решения поставленных проблем, уметь находить полезный дополнительный материал по тематике занятий.
самостоятельная работа	Самостоятельная работа является важнейшим элементом учебного процесса, так как это один из основных методов освоения учебных дисциплин и овладения навыками профессиональной деятельности. На лекциях преподаватель знакомит студентов с основными положениями темы, а дальнейшее усвоение материала связано с самостоятельной работой. Развитие умений самостоятельной работы происходит в процессе подготовки к занятиям. Развиваются умения самостоятельного поиска, отбора и переработки информации. Этому способствуют разные формы постановки заданий для подготовки к занятию ? количество вопросов и их формулировка, указание конкретных источников, разделов, страниц ? или предоставление студентам возможности самостоятельного поиска. Самостоятельная работа с учебниками, учебными пособиями, научной, справочной и популярной литературой, материалами периодических изданий и Интернета, статистическими данными является наиболее эффективным методом получения знаний, позволяет значительно активизировать процесс овладения информацией, способствует более глубокому усвоению изучаемого материала, формирует у студентов свое отношение к конкретной проблеме.
экзамен	Экзамен нацелен на комплексную проверку освоения дисциплины. Экзамен проводится в устной или письменной форме по билетам, в которых содержатся вопросы (задания) по всем темам курса. Обучающемуся даётся время на подготовку. Оценивается владение материалом, его системное освоение, способность применять нужные знания, навыки и умения при анализе проблемных ситуаций и решении практических заданий.

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.

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 01.04.02 "Прикладная математика и информатика" и магистерской программе "Анализ данных и его приложения".

Приложение 2
к рабочей программе дисциплины (модуля)
Б1.В.ОД.5 Дополнительные главы машинного обучения

Перечень литературы, необходимой для освоения дисциплины (модуля)

Направление подготовки: 01.04.02 - Прикладная математика и информатика

Профиль подготовки: Анализ данных и его приложения

Квалификация выпускника: магистр

Форма обучения: очное

Язык обучения: русский

Год начала обучения по образовательной программе: 2017

Основная литература:

Вьюгин, В.В. Математические основы машинного обучения и прогнозирования [Электронный ресурс]: учебное пособие / В.В. Вьюгин. ? Электрон. дан. ? Москва :

МЦНМО, 2014. ? 304 с. ? Режим доступа: <https://e.lanbook.com/book/56397>

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>

Кашина О.А., Миссаров М.Д. Электронный курс 'Анализ данных в среде R', 2014 - Режим доступа:

<https://edu.kpfu.ru/course/view.php?id=833>

Дополнительная литература:

Крянев, А.В. Метрический анализ и обработка данных [Электронный ресурс] / А.В. Крянев, Г.В. Лукин, Д.К. Удумян. - Электрон. дан. - Москва : Физматлит, 2012. - 308 с. -Режим доступа: <https://e.lanbook.com/book/59523>

Компьютерные технологии анализа данных в эконометрике / Д.М. Дайитбегов. - 2-е изд., испр. и доп. - М.:

Вузовский учебник: ИНФРА-М, 2010. - 578 с. - (Научная книга). ISBN 978-5-9558-0191-9 -Режим доступа:

<http://www.znanium.com/bookread.php?book=251791>

Чураков, Е.П. Введение в многомерные статистические методы [Электронный ресурс] : учеб. Пособие-

Санкт-Петербург : Лань, 2016. -148 с. - Режим доступа: <https://e.lanbook.com/book/87598>.

Приложение 3
к рабочей программе дисциплины (модуля)
Б1.В.ОД.5 Дополнительные главы машинного обучения

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

Направление подготовки: 01.04.02 - Прикладная математика и информатика

Профиль подготовки: Анализ данных и его приложения

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

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.