

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
_____ Турилова Е.А.
"___" _____ 20__ г.

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: 2022

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, д.н. (доцент) Миссаров М.Д. (кафедра анализа данных и технологий программирования, отделение фундаментальной информатики и информационных технологий), Moukadas.Missarov@kpfu.ru ; доцент, к.н. (доцент) Шустова Е.П. (кафедра анализа данных и технологий программирования, отделение фундаментальной информатики и информационных технологий), Evgeniya.Shustova@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:

Обучающийся, освоивший дисциплину (модуль):

Должен знать:

метод главных компонент, методы линейного и квадратичного дискриминантного анализа, основы теории машин опорных векторов, методы оценивания и комбинирования классификаторов, методы обучения с подкреплением

Должен уметь:

строить главные компоненты для заданной матрицы данных, классифицировать выборки методами дискриминантного анализа, машин опорных векторов, методов оценивания и комбинирования классификаторов, применять методы обучения с подкреплением.

Должен владеть:

навыками постановки научно-исследовательских задач в различных прикладных областях на основе математических методов машинного обучения.

2. Место дисциплины (модуля) в структуре ОПОП ВО

Данная дисциплина (модуль) включена в раздел "Б1.В.07 Дисциплины (модули)" основной профессиональной образовательной программы 01.04.02 "Прикладная математика и информатика (Анализ данных и его приложения)" и относится к части ОПОП ВО, формируемой участниками образовательных отношений.

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 3 зачетных(ые) единиц(ы) на 108 часа(ов).

Контактная работа - 42 часа(ов), в том числе лекции - 26 часа(ов), практические занятия - 0 часа(ов), лабораторные работы - 16 часа(ов), контроль самостоятельной работы - 0 часа(ов).

Самостоятельная работа - 30 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	Се- местр	Виды и часы контактной работы, их трудоемкость (в часах)						Само- стоя- тель- ная ра- бота
			Лекции, всего	Лекции в эл. форме	Практи- ческие занятия, всего	Практи- ческие в эл. форме	Лабора- торные работы, всего	Лабора- торные в эл. форме	
N	Разделы дисциплины / модуля	Се- местр	Виды и часы контактной работы, их трудоемкость (в часах)						Само- стоя- тель- ная ра- бота
			Лекции, всего	Лекции в эл. форме	Практи- ческие занятия, всего	Практи- ческие в эл. форме	Лабора- торные работы, всего	Лабора- торные в эл. форме	
1.	Тема 1. Методы снижения размерности	3	4	0	0	0	2	0	4
2.	Тема 2. Дискриминантный анализ	3	5	0	0	0	4	0	6
3.	Тема 3. Машины опорных векторов	3	6	0	0	0	2	0	6
4.2	Тема 4. Оценивание и комбинирование классификаторов	3	6	0	0	0	4	0	6
	Тема 5. Обучение с подкреплением	3	5	0	0	0	4	0	8
	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.								30

Тема 2. Дискриминантный анализ

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.

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. Ядерный трюк в кластерном анализе. Ядерные функции для пространства строк. 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. Перекрестная проверка. Бутстреп. Минимизация средних потерь. Разложение смещение-дисперсия. Комбинирование классификаторов (математические мотивировки). Процедуры комбинирования классификаторов. Баггинг. Бустинг. Случайные леса.

Тема 5. Обучение с подкреплением

Примеры задач обучения с подкреплением. Жадные стратегии. Марковские процессы принятия решений. Уравнения Беллмана. Метод временных различий. Q-обучение. Метод Монте-Карло.

Обучение интеллектуального агента хорошей последовательности принятия решений в условиях неполной информации с помощью библиотеки Gym (или Keras) для Python.

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 науки и высшего образования Российской Федерации от 6 апреля 2021 года №245)

Письмо Министерства образования Российской Федерации №14-55-99бин/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)

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

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_\(курс_лекций%2С_К.В.Воронцов\)](http://www.machinelearning.ru/wiki/index.php?title=Машинное_обучение_(курс_лекций%2С_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.
лабораторные работы	Задания на лабораторные работы, пошаговые инструкции по их выполнению и правила оформления отчетов содержатся в Teams: команда "Python в задачах машинного обучения". Лабораторные работы выполняются с использованием реальных данных из открытых репозиториях. Отчёты о выполнении лабораторных работ размещаются студентами в заданиях в указанной выше команде. Преподаватель проверяет и оценивает отчёты.
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.

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

*Приложение 2
к рабочей программе дисциплины (модуля)
Б1.В.07 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: 2022

Основная литература:

1. 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 // Лань : электронно-библиотечная система. - URL: <https://e.lanbook.com/book/56397> (дата обращения: 25.02.2020). - Режим доступа: для авториз. пользователей.
2. Флах П., Машинное обучение. Наука и искусство построения алгоритмов, которые извлекают знания из данных / Флах П. - Москва: ДМК Пресс, 2015. - 400 с. - ISBN 978-5-97060-273-7 - Текст : электронный // ЭБС 'Консультант студента' : [сайт]. - URL : <http://www.studentlibrary.ru/book/ISBN9785970602737.html> (дата обращения: 25.02.2020). - Режим доступа : по подписке.
3. Чураков, Е. П. Введение в многомерные статистические методы : учебное пособие / Е. П. Чураков. - Санкт-Петербург : Лань, 2016. - 148 с. - ISBN 978-5-8114-2149-7. - Текст : электронный // Лань : электронно-библиотечная система. - URL: <https://e.lanbook.com/book/87598> (дата обращения: 25.02.2020). - Режим доступа: для авториз. пользователей.
4. Кашина, О.А. Миссаров М.Д. Электронный курс 'Анализ данных в среде R', 2014. - Текст : электронный. - URL: <https://edu.kpfu.ru/course/view.php?id=833> (дата обращения: 25.02.2020). - Режим доступа: только для студентов и сотрудников КФУ.

Дополнительная литература:

1. Алексеева, Т. В. Информационные аналитические системы: учебник / Т. В. Алексеева, Ю. В. Амириди, В. В. Дик и др.; под ред. В. В. Дика. - Москва : МФПУ Синергия, 2013. - 384 с. - (Университетская серия). - ISBN 978-5-4257-0092-6. - Текст : электронный. - URL: <https://znanium.com/catalog/product/451186> (дата обращения: 25.02.2020). - Режим доступа: по подписке.
2. Ясницкий, Л. Н. Интеллектуальные системы : учебник / Л. Н. Ясницкий. - Москва : Лаборатория знаний, 2016. - 224 с. - ISBN 978-5-00101-417-1. - Текст : электронный // Лань : электронно-библиотечная система. - URL: <https://e.lanbook.com/book/90254> (дата обращения: 25.02.2020). - Режим доступа: для авториз. пользователей.

*Приложение 3
к рабочей программе дисциплины (модуля)
Б1.В.07 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: 2022

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.