

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

Методы оптимизации в задачах анализа данных

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: 2018

Содержание

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

Программу дисциплины разработал(а)(и) доцент, к.н. (доцент) Габидуллина З.Р. (кафедра анализа данных и исследования операций, отделение фундаментальной информатики и информационных технологий), Zulfya.Gabidullina@kpfu.ru

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

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

Шифр компетенции	Расшифровка приобретаемой компетенции
ОПК-1	Способен решать актуальные задачи фундаментальной и прикладной математики
ОПК-2	Способен совершенствовать и реализовывать новые математические методы решения прикладных задач

Обучающийся, освоивший дисциплину (модуль):

Должен знать:

роль и место оптимизации и анализа данных в системе фундаментальных и прикладных математических дисциплин.

Должен уметь:

применять основные результаты методов оптимизации к решению задач анализа данных.

Должен владеть:

основным математическим аппаратом решения оптимизационных задач.

Должен демонстрировать способность и готовность:

формулировать задачи анализа данных в виде оптимизационных математических моделей.

2. Место дисциплины (модуля) в структуре ОПОП ВО

Данная дисциплина (модуль) включена в раздел "Б1.О.03 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 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 4 зачетных(ые) единиц(ы) на 144 часа(ов).

Контактная работа - 32 часа(ов), в том числе лекции - 16 часа(ов), практические занятия - 0 часа(ов), лабораторные работы - 16 часа(ов), контроль самостоятельной работы - 0 часа(ов).

Самостоятельная работа - 76 часа(ов).

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

Форма промежуточного контроля дисциплины: экзамен 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. Основные определения.	2	1	0	1	10
2.	Тема 2. Одномерная идентификация данных и ранжирование.	2	1	0	1	4

N	Разделы дисциплины / модуля	Семестр	Виды и часы контактной работы, их трудоемкость (в часах)			Самостоятельная работа
			Лекции	Практические занятия	Лабораторные работы	
3.	Тема 3. Применение метода главного эксперта в задачах идентификации.	2	1	0	1	2
4.	Тема 4. Различные подходы к построению классификаторов (идентификаторов данных).	2	1	0	3	12
5.	Тема 5. Новый подход (Габидуллиной З.Р.): использование разности Минковского классов данных в постановке оптимизационных задач.	2	2	0	3	12
6.	Тема 6. Математические методы решения оптимизационных задач для построения сепаратора или псевдо-сепаратора классов.	2	3	0	2	10
7.	Тема 7. Метод Гильберта и метод Митчелла-Демьянова-Малоземова для построения классификатора обучающих классов в отделимом случае.	2	3	0	2	11
8.	Тема 8. Универсальные решатели, реализующие критерии линейной отделимости классов (на примере, минимаксной оптимизационной задачи и критерия отделимости Габидуллиной З.Р.).	2	2	0	2	11
9.	Тема 9. Использование метода главного эксперта в медицинской диагностике.	2	2	0	1	4
	Итого		16	0	16	76

4.2 Содержание дисциплины (модуля)

Тема 1. Основные определения.

Типы данных. Математическое представление данных. Классы данных. Типы классов (пересекающиеся и непересекающиеся). Задача классификации. Классификация объектов (четкая и нечеткая, бинарная и мультиклассовая). Пространство признаков. Метрики, используемые для отображения матрицы расстояний между объектами. Алгоритм K- ближайших соседей.

Тема 2. Одномерная идентификация данных и ранжирование.

Идентификация методом разделения. Метод изоляции. Ранжирование параметров.

Тема 3. Применение метода главного эксперта в задачах идентификации.

Постановка задачи идентификации. Основные идеи метода главного эксперта (Демьяновой В.В) для машинного обучения с учителем.

Тема 4. Различные подходы к построению классификаторов (идентификаторов данных).

Классический подход (Bennet K., Mangasarian O.): оптимизационные задачи, применяемые в SVM (Support Vector Machines).

Тема 5. Новый подход (Габидуллиной З.Р.): использование разности Минковского классов данных в постановке оптимизационных задач.

Новый подход (Габидуллиной З.Р.): использование разности Минковского классов данных в постановке оптимизационных задач.

Тема 6. Математические методы решения оптимизационных задач для построения сепаратора или псевдо-сепаратора классов.

Мера толщины сепаратора (отделяющего слоя) в отделимом случае. Мера толщины псевдо-сепаратора (слоя неправильно классифицированных точек) в случае неотделимости классов

Тема 7. Метод Гильберта и метод Митчелла-Демьянова-Малоземова для построения классификатора обучающих классов в отделимом случае.

Метод Гильберта и метод Митчелла-Демьянова-Малоземова для построения классификатора обучающих классов в отделимом случае.

Тема 8. Универсальные решатели, реализующие критерии линейной отделимости классов (на примере, минимаксной оптимизационной задачи и критерия отделимости Габидуллиной З.Р.).

Универсальные решатели, реализующие критерии линейной отделимости классов (на примере, минимаксной оптимизационной задачи и критерия отделимости Габидуллиной З.Р.). Реализация построения классификатора и псевдо-классификатора в пакете Matlab.

Тема 9. Использование метода главного эксперта в медицинской диагностике.

Использование метода главного эксперта для прогнозирования эффективности различных типов терапии для лечения заболеваний. Использование базы данных СНМО-253 (разбиение базы данных на подбазы).

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)

Data Mining Labs - <http://dmlabs.org/>

The R Project - <http://www.r-project.org/>

курс - <http://www.intuit.ru/studies/courses/6/6/info>

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

Вид работ	Методические рекомендации
лекции	Студентам рекомендуется посещать лекционные занятия. На лекционных занятиях, каждому студенту рекомендуется вести подробные записи лекционных материалов, задавать преподавателю возникающие вопросы, перед каждой лекцией просматривать теоретический материал предыдущей лекции для лучшего освоения следующей темы.
лабораторные работы	Для выполнения лабораторной работы требуются навыки работы в пакете прикладных программ Matlab для решения оптимизационных задач, позволяющих строить линейный классификатор отделимых данных, псевдо-классификатор неотделимых данных, вычисления толщины сепаратора и псевдо-сепаратора, построения графической интерпретации полученных результатов. Примеры генерации отделяемых классов данных: 1. $z1=(0,2), z2=(3,3), z3=(2,0)$; $p1=(-2,0), p2=(-3,-3), p3=(0,-2)$. 2. $z1=(-3,3), z2=(3,3), z3=(3,-3)$; $p1=(-1,1), p2=(-2,-2), p3=(3,-3)$. 3. $z1=(-2,3), z2=(3,2), z3=(3,-3)$; $p1=(-1,-1), p2=(0,2), p3=(2,0)$.
самостоятельная работа	Каждому студенту необходимо самостоятельно проработать сценарии решения каждого типа задач, разобранных на занятиях. Следуя рекомендациям преподавателя, выполнять все аудиторные и домашние задания (в том числе, выполнять все лабораторные задания по решению задач отделения данных и готовиться к активному участию в обсуждении теоретических вопросов).
экзамен	Используя лекционный материал, литературу из основного и дополнительного списка подготовиться по следующим вопросам к экзамену: 1. Одномерная идентификация данных и ранжирование. 2. Применение метода главного эксперта в задачах идентификации. 3. Различные подходы к построению классификаторов (идентификаторов данных). 4. Новый подход (Габидуллиной З.Р.): использование разности Минковского классов данных в постановке оптимизационных задач. 5. Математические методы решения оптимизационных задач для построения сепаратора или псевдо-сепаратора классов. 6. Метод Гильберта и метод Митчелла-Демьянова-Малоземова для построения классификатора обучающих классов в отделимом случае. 7. Универсальные решатели, реализующие критерии линейной отделимости классов (на примере, минимаксной оптимизационной задачи и критерия отделимости Габидуллиной З.Р.). 8. Использование метода главного эксперта в медицинской диагностике.

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.О.03 Методы оптимизации в задачах анализа данных

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: 2018

Основная литература:

1. Методы оптимизации: Учебное пособие / А.В. Аттетков, В.С. Зарубин, А.Н. Канатников. - М.: ИЦ РИОР: НИЦ Инфра-М, 2013. - 270 с.: ил.; 60x90 1/16. - (Высшее образование: Бакалавриат). (переплет) ISBN 978-5-369-01037-2, 700 экз.
<http://www.znaniyum.com/bookread.php?book=350985>
2. Компьютерные технологии анализа данных в эконометрике / Д.М. Дайитбегов. - 2-е изд., испр. и доп. - М.: Вузовский учебник: ИНФРА-М, 2010. - 578 с.: 70x100 1/16. - (Научная книга). (переплет) ISBN 978-5-9558-0191-9
<http://www.znaniyum.com/bookread.php?book=251791>
3. Ашманов С.А. Тимохов А.В. Теория оптимизации в задачах и упражнениях. - Санкт-Петербург: Лань, 2012. - 448с.
http://e.lanbook.com/books/element.php?pl1_id=3799
4. Лесин В. В., Лисовец Ю. П. Основы методов оптимизации. - Санкт-Петербург: Лань, 2011. - 352 с.
http://e.lanbook.com/books/element.php?pl1_id=1552
5. Интеллектуальный анализ временных рядов [Электронный ресурс]: Учебное пособие / Н.Г. Ярушкина, Т.В. Афанасьева, И.Г. Перфильева. - М.: ИД ФОРУМ: ИНФРА-М, 2012. - 160 с. - Режим доступа: <http://www.znaniyum.com/bookread.php?book=249314>

Дополнительная литература:

1. Степанов, Роман Григорьевич. Технология Data Mining: Интеллектуальный анализ данных: учебное пособие / Р. Г. Степанов; Казан. гос. ун-т. Казань: Казанский государственный университет, 2009. 110 с
2. Барсебян, А. А. Анализ данных и процессов: учеб. пособие / А. А. Барсебян, М. С. Куприянов, И. И. Холод, М. Д. Тесс, С. И. Елизаров. ? 3-е изд., перераб. и доп. ? СПб.: БХВ-Петербург, 2009. ? 512 с.: ил. + CD-ROM ? (Учебная литература для вузов).
<http://www.znaniyum.com/bookread.php?book=350638>
3. Заботин, И.Я. Методы и вычислительные приемы в линейном программировании [Электронный ресурс] : учебное пособие / И.Я. Заботин, Я.И. Заботин. ? Электрон. дан. ? Казань : КФУ (Казанский (Приволжский) федеральный университет), 2014. ? 116 с. ? Режим доступа: http://e.lanbook.com/books/element.php?pl1_id=72810 ? Загл. с экрана.
4. 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://www.znaniyum.com/bookread.php?book=451186>

Приложение 3
к рабочей программе дисциплины (модуля)
Б1.О.03 Методы оптимизации в задачах анализа данных

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: 2018

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

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

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

Браузер Mozilla Firefox

Браузер Google Chrome

Adobe Reader XI или Adobe Acrobat Reader DC

Kaspersky Endpoint Security для Windows

Учебно-методическая литература для данной дисциплины имеется в наличии в электронно-библиотечной системе "ZNANIUM.COM", доступ к которой предоставлен обучающимся. ЭБС "ZNANIUM.COM" содержит произведения крупнейших российских учёных, руководителей государственных органов, преподавателей ведущих вузов страны, высококвалифицированных специалистов в различных сферах бизнеса. Фонд библиотеки сформирован с учетом всех изменений образовательных стандартов и включает учебники, учебные пособия, учебно-методические комплексы, монографии, авторефераты, диссертации, энциклопедии, словари и справочники, законодательно-нормативные документы, специальные периодические издания и издания, выпускаемые издательствами вузов. В настоящее время ЭБС ZNANIUM.COM соответствует всем требованиям федеральных государственных образовательных стандартов высшего образования (ФГОС ВО) нового поколения.

Учебно-методическая литература для данной дисциплины имеется в наличии в электронно-библиотечной системе Издательства "Лань", доступ к которой предоставлен обучающимся. ЭБС Издательства "Лань" включает в себя электронные версии книг издательства "Лань" и других ведущих издательств учебной литературы, а также электронные версии периодических изданий по естественным, техническим и гуманитарным наукам. ЭБС Издательства "Лань" обеспечивает доступ к научной, учебной литературе и научным периодическим изданиям по максимальному количеству профильных направлений с соблюдением всех авторских и смежных прав.