

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное автономное образовательное учреждение высшего образования
"Казанский (Приволжский) федеральный университет"
Высшая школа бизнеса КФУ



подписано электронно-цифровой подписью

Программа дисциплины Аналитические системы (Big Data)

Направление подготовки: 38.04.02 - Менеджмент

Профиль подготовки: IT-менеджмент

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) Кох И.А. (Кафедра финансовых рынков и финансовых институтов, Институт управления, экономики и финансов), IAKoh@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-2	Способен применять современные техники и методики сбора данных, продвинутые методы их обработки и анализа, в том числе использовать интеллектуальные информационно-аналитические системы, при решении управленческих и исследовательских задач

Обучающийся, освоивший дисциплину (модуль):

Должен знать:

Базовые принципы сбора и предварительной обработки больших данных (Big Data)

- Подходы к разведочному анализу данных
- Основные технологии анализа больших данных (Big Data)

Должен уметь:

Разрабатывать программы загрузки наборов больших данных

- Анализировать наборы больших данных
- Строить модели обработки больших данных

Должен владеть:

- Терминологией курса
- Современными технологиями разведочного анализа данных
- Подходами к анализу данных: (регрессия, прогнозирование, кластеризация, классификация, нейронные сети)

Современными пакетами обработки больших данных

Должен демонстрировать способность и готовность:

самостоятельно применять в определенном контексте полученные в процессе изучения курса 'Аналитические системы Big Data' знания, умения и навыки.

2. Место дисциплины (модуля) в структуре ОПОП ВО

Данная дисциплина (модуль) включена в раздел "Б1.В.ДВ.07.01 Дисциплины (модули)" основной профессиональной образовательной программы 38.04.02 "Менеджмент (IT-менеджмент)" и относится к дисциплинам по выбору части ОПОП ВО, формируемой участниками образовательных отношений.

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 - 26 часа(ов), в том числе лекции - 8 часа(ов), практические занятия - 18 часа(ов), лабораторные работы - 0 часа(ов), контроль самостоятельной работы - 0 часа(ов).

Самостоятельная работа - 10 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	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	Лабораторные работы, всего	Лабораторные в эл. форме	
1.	Тема 1. Тема 1. Определение больших данных. Технологии хранения больших данных.	2	2	0	2	0	0	0	2
2.	Тема 2. Тема 2. Процесс анализа больших данных. CRISP-DM. Технологии анализа больших данных. Научные проблемы в области больших данных.	2	2	0	6	0	0	0	2
3.	Тема 3. Тема 3. Методы анализа больших данных.	2	2	0	4	0	0	0	3
4.	Тема 4. Тема 4. Программы статистической обработки информации. Представление возможностей пакета R для целей анализа больших данных.	2	2	0	6	0	0	0	3
	Итого		8	0	18	0	0	0	10

4.2 Содержание дисциплины (модуля)

Тема 1. Тема 1. Определение больших данных. Технологии хранения больших данных.

Большие данные (big data) в информационных технологиях - совокупность подходов, инструментов и методов обработки структурированных и неструктурированных данных огромных объёмов и значительного многообразия для получения воспринимаемых человеком результатов, эффективных в условиях непрерывного прироста, распределения по многочисленным узлам вычислительной сети, сформировавшихся в конце 2000-х годов, альтернативных традиционным системам управления базами данных и решениям класса Business Intelligence. В данную серию включают средства массово-параллельной обработки неопределённо структурированных данных, прежде всего, решениями категории NoSQL, алгоритмами MapReduce, программными каркасами и библиотеками проекта Hadoop.

В качестве определяющих характеристик для больших данных отмечают "три V?" объём (англ. volume, в смысле величины физического объёма), скорость (англ. velocity в смыслах как скорости прироста, так и необходимости высокоскоростной обработки и получения результатов), многообразие (англ. variety, в смысле возможности одновременной обработки различных типов структурированных и полуструктурированных данных).

Тема 2. Тема 2. Процесс анализа больших данных. CRISP-DM. Технологии анализа больших данных. Научные проблемы в области больших данных.

Методы и техники анализа, применимые к большим данным:

методы класса Data Mining: обучение ассоциативным правилам (англ. association rule learning), классификация (методы категоризации новых данных на основе принципов, ранее применённых к уже наличествующим данным), кластерный анализ, регрессионный анализ;

краудсорсинг - категоризация и обогащение данных силами широкого, неопределённого круга лиц, привлечённых на основании публичной оферты, без вступления в трудовые отношения;

смешение и интеграция данных (англ. data fusion and integration) - набор техник, позволяющих интегрировать разнородные данные из разнообразных источников для возможности глубинного анализа, в качестве примеров таких техник, составляющих этот класс методов приводятся цифровая обработка сигналов и обработка естественного языка (включая тональный анализ);

машинное обучение, включая обучение с учителем и без учителя, а также Ensemble learning (англ.) - использование моделей, построенных на базе статистического анализа или машинного обучения для получения комплексных прогнозов на основе базовых моделей (англ. constituent models, ср. со статистическим ансамблем в статистической механике);

искусственные нейронные сети, сетевой анализ, оптимизация, в том числе генетические алгоритмы;

распознавание образов;

прогнозная аналитика;

имитационное моделирование;

пространственный анализ (англ. Spatial analysis) - класс методов, использующих топологическую, геометрическую и географическую информацию в данных;

статистический анализ, в качестве примеров методов приводятся А/В-тестирование и анализ временных рядов; визуализация аналитических данных - представление информации в виде рисунков, диаграмм, с использованием интерактивных возможностей и анимации как для получения результатов, так и для использования в качестве исходных данных для дальнейшего анализа.

Тема 3. Методы анализа больших данных.

Понятие прогноза и предвидения. Отличие прогнозирования от предвидения. Закон распределения случайной величины. Статистические оценки параметров. Доверительные области. Теория моментов. Корреляционный анализ. Использование модели множественной линейной регрессии для прогнозирования экономических показателей. Доверительные интервалы для зависимой переменной. Сглаживание временных рядов. Динамические модели с распределенными лагами. Стационарные временные ряды. Тестирование стационарности. Коинтеграция. Анализ временных рядов. Адаптивные и мультипликативные методы прогнозирования. Экспоненциальное сглаживание. Авторегрессионные модели. Модели скользящего среднего. Интегрированные процессы. Идентификация авторегрессионной модели скользящего среднего. Прогнозирование с моделями временных рядов. Доверительные интервалы прогноза. Дисперсионный анализ влияния качественных факторов. Ранговые методы. Факторный анализ. Метод главных факторов. Многомерное шкалирование. Классическая модель многомерного шкалирования. Неметрические методы. Кластерный анализ. Дискриминантный анализ. Многомерный статистический анализ.

Тема 4. Программы статистической обработки информации. Представление возможностей пакета R для целей анализа больших данных.

R -- статистический пакет для обработки больших данных, один из лидеров рынка в области некоммерческих статистических продуктов, предназначенных для проведения прикладных исследований в различных предметных областях. Применение программы для решения прикладных задач прогнозирования: ввод и хранение данных; возможность использования переменных разных типов; частотность признаков, таблицы, графики, таблицы сопряженности, диаграммы; первичная описательная статистика; маркетинговые и медиа исследования; анализ данных маркетинговых и медиа исследований.

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 планируемому результату 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)

Дистрибутив IDE RStudio - <https://rstudio.com/>

Дистрибутив пакета R The R Project for Statistical Computing - <https://www.r-project.org/>

Официальный сайт архива пакетов для R - <https://cran.r-project.org/>

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

Вид работ	Методические рекомендации
лекции	Каждая тема содержит лекционный материал, список литературы для самостоятельного изучения, вопросы и задания для подготовки к семинарским и/или практическим занятиям, а также материалы для самостоятельной работы. Необходимо заранее обеспечить себя этими материалами и литературой или доступом к ним.
практические занятия	Семинар по дисциплине является аудиторным занятием, в процессе которого преимущественно осуществляется контроль знаний, полученных студентом самостоятельно. В связи с этим такое занятие начинается либо с устного опроса либо с контрольной работы, которая может проводиться по лекционному материалу темы, литературным источникам, указанным по данной теме заданиям для самостоятельной работы.
самостоятельная работа	Лекционный материал и указанные литературные источники по соответствующей теме необходимо изучить до посещения соответствующего лекционного занятия, так как лекция в аудитории предполагает раскрытие актуальных и проблемных вопросов рассматриваемой темы, а не содержания лекционного материала. Таким образом, для понимания того, что будет сказано на лекции, необходимо получить базовые знания по теме, которые содержатся в лекционном материале.
экзамен	Подготовка к экзамену является заключительным этапом изучения дисциплины и является средством текущего контроля. В процессе подготовки к экзамену выявляются вопросы, по которым нет уверенности в ответе либо ответ студенту не ясен. В связи с этим подготовка к экзамену заключается в том, чтобы изучить лекционный материал и указанные по теме литературные источники выполнить задания для самостоятельной работы.

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 38.04.02 "Менеджмент" и магистерской программе "IT-менеджмент".

Приложение 2
к рабочей программе дисциплины (модуля)
Б1.В.ДВ.07.01 Аналитические системы (Big Data)

Перечень литературы, необходимой для освоения дисциплины (модуля)

Направление подготовки: 38.04.02 - Менеджмент

Профиль подготовки: IT-менеджмент

Квалификация выпускника: магистр

Форма обучения: очное

Язык обучения: русский

Год начала обучения по образовательной программе: 2022

Основная литература:

1. Павлов, В. М. Искусство решать сложные задачи: системный подход : монография / В. М. Павлов. - 5-е изд. - Москва : Издательско-торговая корпорация 'Дашков и К-', 2019. - 184 с. - ISBN 978-5-394-03425-1. - Текст : электронный. - URL: <https://znanium.com/catalog/product/1091820> (дата обращения: 19.04.2021)
2. Орлова, И. В. Экономико-математическое моделирование: Практическое пособие по решению задач / И.В. Орлова. - 2-е изд., испр. и доп. - Москва : Вузовский учебник: НИЦ Инфра-М, 2012. - 140 с. ISBN 978-5-9558-0107-0. - Текст : электронный. - URL: <https://znanium.com/catalog/product/359462> (дата обращения: 19.04.2021)
3. Непомнящая, Н. В. Статистика: общая теория статистики, экономическая статистика. Практикум/ Н.В.Непомнящая, Е.Г.Григорьева. - Краснояр.: СФУ, 2015. - 376 с.: ISBN 978-5-7638-3185-6. - Текст : электронный. - URL: <https://znanium.com/catalog/product/549841> (дата обращения: 19.04.2021).

Дополнительная литература:

1. Прогнозирование и планирование в условиях рынка: Учебное пособие / Т.Н. Бабич, И.А. Козьева, Ю.В. Вертакова, Э.Н. Кузьбожев. - Москва : НИЦ ИНФРА-М, 2013. - 336 с. (Высшее образование: Бакалавриат). ISBN 978-5-16-004577-1. - Текст : электронный. - URL: <https://znanium.com/catalog/product/407700> (дата обращения: 19.04.2021)
2. Дадян, Э. Г. Методы, модели, средства хранения и обработки данных : учебник / Э.Г. Дадян, Ю.А. Зеленков. - Москва : Вузовский учебник : ИНФРА-М, 2017. - 168 с. - ISBN 978-5-9558-0490-3. - Текст : электронный. - URL: <https://znanium.com/catalog/product/543943> (дата обращения: 19.04.2021)
3. Агальцов В.П. Базы данных. В 2-х кн. Книга 2. Распределенные и удаленные базы данных : учебник / В.П. Агальцов. - Москва : ИД 'ФОРУМ' : ИНФРА-М, 2017. - 271 с. : ил. - (Высшее образование). - ISBN 978-5-8199-0394-0. - Текст : электронный. - URL: <https://znanium.com/catalog/product/652917> (дата обращения: 19.04.2021)
4. Козлов, А. Ю. Статистический анализ данных в MS Excel: Учебное пособие / А.Ю. Козлов, В.С. Мхитарян, В.Ф. Шишов. - Москва : ИНФРА-М, 2014. - 320 с. (Высшее образование). ISBN 978-5-16-004579-5. - Текст : электронный. - URL: <https://znanium.com/catalog/product/429722> (дата обращения: 19.04.2021)

*Приложение 3
к рабочей программе дисциплины (модуля)
Б1.В.ДВ.07.01 Аналитические системы (Big Data)*

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

Направление подготовки: 38.04.02 - Менеджмент

Профиль подготовки: IT-менеджмент

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.