

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РФ
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Дагестанский государственный университет»

Факультет информатики и информационных технологий

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

Современные методы обработки больших данных

Кафедра информационных технологий и БКС

Образовательная программа

09.04.02 Информационные системы и технологии

Профиль подготовки:

Информационно-телекоммуникационные системы и сети

Уровень высшего образования:

магистратура

Форма обучения:

очно-заочная

Статус дисциплины:

входит в обязательную часть ОПОП

Рабочая программа дисциплины «Современные методы обработки больших данных» составлена в соответствии с требованиями Федерального государственного образовательного стандарта высшего образования – магистратура по направлению подготовки 09.04.02 Информационные системы и технологии от «19» сентября 2017 г. № 917.

Составитель:



Ахмедова З.Х., доцент каф. ИТиБКС

Рабочая программа одобрена на заседании кафедры «ИТиБКС. Протокол № 11 от 28.06 2021 г

Зав кафедрой ИТиБКС  З.Х. Ахмедова

Одобрена на заседании Методической комиссии факультета Информатики и информационных технологий

Протокол № 4 от 28.06 2021 г

Председатель  Бакмаев А.Ш.

Рабочая программа согласована с учебно-методическим управлением

907 2021г 

Аннотация рабочей программы дисциплины

Дисциплина «Современные методы обработки больших данных» входит в обязательную часть базового модуля образовательной программы магистратуры по направлению подготовки 09.04.02-Информационные системы и технологии.

Дисциплина реализуется на факультете информатики и информационных технологий кафедрой ИТиБКС.

Основная цель данной дисциплины – формирование у студентов базовых теоретических знаний в области теории вероятностей и математической статистики и способности к применению технологий обработки данных (в том числе big data) и машинного обучения к решению прикладных задач, связанных с оказанием финансовых услуг.

Преподавание дисциплины предусматривает проведение следующих видов учебных занятий: лекции, лабораторные работы и самостоятельную работу.

Дисциплина нацелена на формирование следующих компетенций выпускника: ОПК-1, ОПК-4, Профессиональных – ПК-1

Рабочая программа дисциплины предусматривает проведение следующих видов контроля успеваемости в форме контрольных работ и промежуточный контроль в форме зачета

Объем дисциплины 5 зачетных единиц, в том числе в академических часах по видам учебных занятий:

Се- мestr	Учебные занятия							СРС, в том числе экза- мен	Форма проме- жуточной атте- стации (зачет, дифференциро- ванный зачет, экзамен)
	в том числе								
	Контактная работа обучающихся с преподавателем								
	Все го	из них							
Всег о		Лекции	Лабора- торные заняти я	Практи- ческие занятия	Кон троль	консу ль- тации			
3	180	24	8	16				156	зачет

1. Цель освоения дисциплины

Курс «Современные методы обработки больших данных» имеет своей целью: формирование у студентов профессиональной компетенции в области разработки и использования систем обработки и анализа больших массивов данных. Данная цель соотносится с целью образовательной программой в частности с технологией разработки специализированных программных систем, отвечающих за обработку больших данных. Изучение данной дисциплины готовит выпускника к выполнению следующих профессиональных задач:

- Постановка задачи анализа данных.
- Предварительная обработка данных.
- Визуализация данных.
- Разработка, реализация и применение методов интеллектуального анализа данных к большим массивам данных.
- Представление результатов работы.

Задачи дисциплины

Основные задачи освоения дисциплины:

Студент должен **знать** методы анализа и хранения больших объемов данных, этапы жизненного цикла обработки больших данных, языки, наиболее приспособленные для обработки и аналитики больших данных, способы организации хранения и доступа к большим данным; **уметь** выполнять элементы анализа данных и интерпретировать результаты, различать характеристики SQL и NoSql БД, формулировать алгоритмы в парадигме MapReduce, выбрать подходящий инструмент анализа больших данных, выбрать подходящую технологию хранения больших данных.; **владеть** математическими методами анализа данных, языками и компьютерными методами обработки.

2. Место дисциплины (модуля) в структуре образовательной программы магистратуры

Курс «Современные методы обработки больших данных» относится к базовой части блока Б1 дисциплин Дисциплины (модули).

Для изучения дисциплины студент должен владеть знаниями, умениями и навыками по дисциплинам:

Дискретная математика, Алгебраические структуры, Основы программирования, Алгоритмы вычислительной математики, Конструирование алгоритмов и структур данных, Теория алгоритмов и вычислительных процессов, Основы теории вероятностей и статистических методов, Алгоритмы и структуры данных, Математическая логика и теория алгоритмов, Интеллектуальный анализ данных.

Знания, получаемые при изучении дисциплины «Современные методы обработки больших данных» используются при изучении профессиональных дисциплин Распределенные задачи и алгоритмы, Программирование в компьютерных сетях, Облачные вычисления, Мультиагентные системы, а также для работ над дипломной и магистерской работой.

3. Компетенции обучающегося, формируемые в результате освоения дисциплины.

Процесс изучения дисциплины направлен на формирование элементов следующих компетенций в соответствии с ФГОС ВО по данному направлению:

Изучение данной учебной дисциплины направлено на формирование у обучающихся следующими **профессиональными компетенциями:**

Код и наименование компетенции из ОПОП	Код и наименование индикатора достижения компетенций	Планируемые результаты обучения	Процедура освоения
ОПК-1. Способен самостоятельно приобретать, развивать и применять математические, естественнонаучные, социально-экономические и профессиональные знания для решения нестандартных задач, в том числе в новой или незнакомой среде и в междисциплинарном контексте;	ИД-1.1. Знает основы математики, физики, вычислительной техники и программирования.	Знает основы математики, физики, вычислительной техники и программирования.	Устный опрос
	ИД-1.2. Умеет решать стандартные профессиональные задачи с применением естественнонаучных и общеинженерных знаний, методов математического анализа и моделирования.	Умеет решать стандартные профессиональные задачи с применением естественнонаучных и общеинженерных знаний, методов математического анализа и моделирования.	
	ИД-1.3. Имеет навыки теоретического и экспериментального исследования объектов профессиональной деятельности.	Имеет навыки теоретического и экспериментального исследования объектов профессиональной деятельности	
ОПК-4. Способен применять на практике новые научные принципы и методы исследований.	ИД-4.1. Знает основные стандарты оформления технической документации на различных стадиях жизненного цикла информационной системы.	Знает основные стандарты оформления технической документации на различных стадиях жизненного цикла информационной системы	Круглый стол
	ИД-4.2. Умеет применять стандарты оформления технической документации на различных стадиях	Умеет применять стандарты оформления технической документации на различных стадиях	

	жизненного цикла информационной системы.	жизненного цикла информационной системы.	
	ИД-4.3. Имеет навыки составления технической документации на различных этапах жизненного цикла информационной системы.	Имеет навыки составления технической документации на различных этапах жизненного цикла информационной системы	
ПК-1 Способен разрабатывать и исследовать модели объектов профессиональной деятельности, предлагать и адаптировать методики, определять качество проводимых исследований, составлять отчеты о проделанной работе, обзоры, готовить публикации	ИД1.1 ПК-1.1 Знает отечественную и международную нормативную базу в области профессиональной деятельности, актуальную научную проблематику в области информационных систем и технологий, методы, средства и практику планирования, организации, проведения и внедрения научных исследований, методы разработки информационных моделей хозяйствующих субъектов, методы формирования показателей эффективности и конкурентоспособности научно-исследовательских работ в области информационных систем и технологий, лучшие практики отечественного и зарубежного опыта разработки и исследований моделей объектов профессиональной деятельности	Знает отечественную и международную нормативную базу в области профессиональной деятельности, актуальную научную проблематику в области информационных систем и технологий, методы, средства и практику планирования, организации, проведения и внедрения научных исследований	КОЛЛОКВИУМ
	ИД 1.2. Умеет применять актуальную нормативную документацию в области профессиональной	Умеет разрабатывать и исследовать способы теоретических и экспериментальных моделей объектов	

	<p>деятельности, анализировать новую научную проблематику и научно-исследовательские разработки в области информационных систем и технологий, применять методы и средства планирования, организации, проведения и внедрения научных исследований, применять методы разработки информационных, объектных, документных моделей хозяйствующих субъектов, проектировать систему управления научно-исследовательскими работами в организации, готовить научные и научно-практические публикации в области профессиональной деятельности</p> <p>ИД1.3.ПК-1.3 Владеет навыками проведения анализа новых направлений исследований в области профессиональной деятельности, обоснования перспектив проведения исследований в области профессиональной деятельности, формирования программ проведения исследований в новых направлениях, осуществления методического руководства проведения научных исследований рабочими группами, анализа результатов работ соисполнителей, участвующих в выполнении работ с другими организациями</p>	<p>профессиональной деятельности</p> <p>Владеет навыками проведения анализа новых направлений исследований в области профессиональной деятельности, обоснования перспектив проведения исследований в области профессиональной деятельности</p>	
--	--	--	--

4. Объем, структура и содержание дисциплины.

4.1 Объем дисциплины

Объем дисциплины составляет 5 зачетных единиц, 180 часов.

4.2. Структура дисциплины

Разделы дисциплины			Трудоемкость освоения раздела дисциплины, час				Формы текущего контроля успеваемости Форма промежуточной аттестации
Наименование разделов	семестр	Номер модуля, в который включен раздел	Лекции	Практические занятия	Лабораторные работы	Самостоятельная работа	
Модуль 1.							
Обзор Big-Data. Методы и средства. Используемые программы. Особенности	1	1	2			20	Опрос
SAS Interprise Miner. Введение. Возможности. Инструменты	1	1			2	12	Опрос
Итого по модулю:			2		2	32	Контрольная работа
Модуль 2.							
Создание проекта. Определение источника данных. Исследование источника данных	1	2	2		2	12	Опрос
Прогнозное моделирование	1	2			2	14	Опрос
Итого по модулю:			2		4	28	Контрольная работа
Модуль 3.							
Прогнозная модель, использующая дерево решений	1	3	2			12	Опрос
Прогнозное моделирование: работа с регрессионными моделями	1	3			2	18	Опрос
Итого по модулю:			2		2	32	Контрольная работа

Модуль 4							
Прогнозное моделирование: нейронные сети			2		2	12	Опрос
Кластерный анализ.					2	18	Контрольная работа
Итого по модулю:			2		4	30	Контрольная работа
Модуль 5							
Анализ потребительской корзины					2	20	Круглый стол
Оценка моделей. Сравнение моделей					2	14	Контрольная работа
Итого по модулю:					4	32	Опрос
Итого по дисциплине	180		8		16	156	

4.1. Содержание дисциплины

4.1.1. Занятия лекционного типа

№ раздела	Наименование раздела	Содержание раздела	Форма текущего контроля	Разработ. с участием представителей работодателей
1	2	3	4	5
1	Введение в большие данные. Понятие Data Minig. Прикладные инструменты для работы с Big Data. Технология MapRaduce. Hadoop.	Предпосылки формирования тренда больших данных <ul style="list-style-type: none"> ▪ Основные вызовы больших данных (4V) ▪ Определение термина "большие данные" ▪ Базовое представление о Map Reduce и Hadoop ▪ Представление о работе аналитика Инструменты для обработки больших данных <ul style="list-style-type: none"> ▪ Знакомство с языками и прикладными пакетами для обработки больших данных. ▪ Рассмотрение общей концепции и синтаксиса языка R (примеры). 	ЛР	
2	Технологии анализа данных: Жизненный цикл анализа больших данных, стандарты. Когнитивный анализ данных Визуализация больших данных.	Аналитика больших данных. <ul style="list-style-type: none"> ▪ Процесс аналитики ▪ Стандарты жизненного цикла Big Data: CRISP-DM ▪ Принципы и инструменты аналитики ▪ Задачи и компетенции аналитиков Big Data ▪ Big Data как рынок ▪ Стек технологий 	ЛР РГЗ	

		<ul style="list-style-type: none"> ▪ По поддержки принятия решений ▪ Игроки на рынке BD ▪ Крупнейшие проекты BD в России <p>Когнитивный анализ данных</p> <ul style="list-style-type: none"> ▪ Введение в Data Mining – понятие, структура, составляющие и сопутствующие науки. ▪ Задачи Data Mining и способы их решения. Классификация методов DM. ▪ Области применения DM. ▪ Классы систем DM. ▪ Процесс накопления и анализа данных: Азбука когнитивного анализа. <p>Аналитика больших данных.</p> <p>Математическая статистика</p> <p>Основные понятия статистики и дескриптивный анализ</p> <ul style="list-style-type: none"> ▪ Шкалы измерений. ▪ Генеральная совокупность и выборка. ▪ Нормальное распределение. Уровень статистической достоверности. ▪ Свойства описательных статистик (Дескриптивный анализ) ▪ Визуальное представление данных ▪ Меры изменчивости <p>Методы DATA MINING</p> <ul style="list-style-type: none"> ▪ Данные & знания ▪ Типовые задачи Data Mining ▪ Обучаемые и необучаемые задачи ▪ Жизненный цикл проекта DM ▪ Математический аппарат DM ▪ Стандарты DM <p>Методы анализа на графах</p> <p>Случайные графы, безмасштабные графы, социальные сети – сети тесного мира. Закономерности, методы кластеризации на графах.</p>		
		<p>Прикладные инструменты анализа данных. Корреляция.</p> <ul style="list-style-type: none"> ▪ Готовые комплексные решения: Weka, RapidMiner, Knime, Orange IBM SPSS Modeler (в прошлом Clementine) ▪ Инструменты визуализации: <i>Tableau</i>, Фреймворки на JS, D3 ▪ Корреляция <ul style="list-style-type: none"> ○ Понятие корреляции ○ Значимость коэффициента корреляции ○ Виды связи между переменными 		

3	Технологии хранения больших данных. Распределенные хранилища, NoSql хранилища, классификация и примеры.	Хранилища данных. Регрессия <ul style="list-style-type: none"> ▪ Хранилища данных <ul style="list-style-type: none"> ○ OLAP и OLTP системы ○ Характеристики BigData и хранилища данных ○ Почему не реляционные СУБД? <ul style="list-style-type: none"> ▪ Требования к хранилищам данных ▪ Регрессионный анализ Распределенные базы данных NoSQL. Решение задач Data Mining. Задачи классификации, кластеризации <ol style="list-style-type: none"> 1. Распределенные базы данных NoSQL <ul style="list-style-type: none"> ▪ Типы NoSQL ▪ Репликация и шардинг ▪ Пример NoSQL БД Задачи классификации и кластеризации <ul style="list-style-type: none"> ▪ Desision Tree ▪ RandomForest ▪ K-means ▪ R и MapReduce Распределенные базы данных NoSQL. Прмеры: HBase, Cassandra, Neo4j, MongoDB. Распределенные файловые системы (РФС). Структура РФС, Требования к РФС, Примеры: HDFS, Google, LustreFS	ЛР Д	
---	---	--	---------	--

Занятия семинарского типа – не предусмотрены.

4.1.2. Лабораторные занятия

№ работы	№ раздела дисциплины	Наименование лабораторных работ
1	1	Ознакомление с синтаксисом языка R для анализа данных.
2	1	Способы подготовки и отображения данных в R . Возможности ввода/вывода.
3	2	Решение задач на больших графах .
4	2	Способы анализа данных в R. Получение первичных элементарных характеристик о наборах данных (элементарные статистики). Способы импорта/экспорта данных.
5	2	Работа с диаграммами и графиками в R .
6	2	Проверка статистических гипотез
7	2	Корреляционный анализ и регрессионный анализ данных
8	2	Решение задач Data Mining. Задачи классификации, кластеризации: деревья решений, RandomForest, k-means.
9	3	Изучение принципов работы распределенных баз данных
10	3	Развертывание локального кластера Hadoop. Подсчёт слов в тексте, с помощью MapReduce.
11	1-3	Круглый стол: Совместное обсуждение результатов РГЗ
12	1-3	Обсуждение итогов курса

5. Образовательные технологии

При проведении занятий рекомендуется использование активных и интерактивных форм занятий (компьютерных симуляций, проектных методик, мозгового штурма, разбора конкретных ситуаций, иных форм) в сочетании с внеаудиторной работой с целью формирования и развития профессиональных навыков обучающихся.

6. Учебно-методическое обеспечение самостоятельной работы студентов

№	Вид СРС	Перечень учебно-методического обеспечения дисциплины по выполнению самостоятельной работы
1	2	3
1	Раздел 1. Чтение публикаций по истории развития Big Data, Data Mining [3-5]. Изучение языка R и Python [7,8] осн. список, [1-4] - дополнительный.	Приходько Т.А. Методические указания по выполнению лабораторных работ по дисциплине «Обработка больших данных», утвержденные кафедрой вычислительных технологий
2	Раздел 2. Изучение части курса Introduction to Data Science [6], посвященной визуализации. Визуализация стандартных наборов данных при помощи Tableau.	Приходько Т.А. Методические указания по выполнению лабораторных работ по дисциплине «Обработка больших данных», утвержденные кафедрой вычислительных технологий.
3	Раздел 3. Изучение парадигмы Map Reduce. Подсчет кол-ва слов, реализация алгоритма k-means в рамках парадигмы Map Reduce с использованием Hadoop. Чтение публикаций о распределенных хранилищах данных, их особенностях и принципах построения распределенных файловых систем.	Источники основной и дополнительной литературы

Для лиц с ограниченными возможностями здоровья предусмотрена организация консультаций с использованием электронной почты.

1. Фонд оценочных средств для проведения текущего контроля успеваемости, промежуточной аттестации по итогам освоения дисциплины.

7.1. Типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций в процессе освоения образовательной программы

Примеры РГЗ – задания на анализ данных

Подобрать данные для таблицы, приведенной ниже и проанализировать их взаимное влияние, отобразить корреляцию:

- Роста ВВП на прирост населения
- Прироста населения на динамику безработицы
- Прирост людей с высшим образованием на рост промышленного производства
- Прирост людей с высшим образованием на развитие науки
- Прирост людей с высшим образованием на динамику доходов на душу населения
- Динамику безработицы на динамику преступности
- С помощью регрессионного анализа найдите зависимые переменные и

поясните влияние на них независимых переменных.

- h. С помощью функции predict() (см. лекции и help()) постройте прогноз по столбцу, соответствующему варианту.

Годы	Численность населения	Рост ВВП	Динамика безработицы	Динамика промышленного производства	Прирост людей, получивших очное высшее образование	Развитие науки (высокотехнологичных отраслей)	Динамика доходов на душу населения	Динамика преступности
	1	2	3	4	5	6	7	8
01.01.1990								
...								
01.01.2015								

Отчет по выполнению РГЗ должен содержать:

2. постановку задачи;
3. сформированный набор данных;
4. тексты скриптов на языке R;
5. результаты тестов на проверку гипотез о корреляции, оценка регрессии, вычисление корреляции в текстовом и графическом виде.
6. ясное и подробное пояснение каждого результата, словесную трактовку графиков;
7. выводы;
8. список использованной литературы.

Пример задания для контрольной работы:

1. Перечислите стандарты жизненного цикла больших данных.
2. Перечислите и охарактеризуйте методы интеллектуального анализа данных.
3. Перечислите классы хранилищ больших данных, назовите различия между ними.

Перечень вопросов, для подготовки к зачету

1. Определение больших данных, ключевые характеристики. Примеры задач больших данных. Основные виды данных.
2. Дать краткую сравнительную характеристику инструментария ПО для анализа данных.
3. Охарактеризовать конструкции языка R Перечислить типы языка R, привести примеры.
4. Роль аналитика по данным (Data Scientist). Ключевые компетенции аналитика. Отличия BI от Data Science.
5. «Жизненный цикл» проекта по аналитике больших данных. Типовая архитектура проекта в области больших данных. Перечислить используемые технологии, указать степень вовлеченности каждой из технологий на каждом этапе работы над проектом. Перечислить основные роли исполнителей

- проекта.
6. Что такое Data Mining? Основные задачи и методы Data Mining. Этапы интеллектуального анализа данных. Методы интеллектуального анализа данных.
 7. Что такое ИИ? Декатлон?
 8. Роль гипотез в процессе познания. Какие факторы используются для уточнения гипотез?
 9. Основные понятия статистики и дескриптивный анализ:
 10. Шкалы измерений. Генеральная совокупность и выборка. Нормальное распределение. Уровень статистической достоверности.
 11. Корреляция и регрессионный анализ. Коэффициент корреляции. Графическое представление. Постановка задачи регрессионного анализа.
 12. Пояснить термин "Линейная регрессия". Привести примеры использования регрессионного анализа.
 13. Классификация и кластеризация – суть и назначение. Метрики. Постановка задачи кластеризации. Методы кластеризации на графах. Отличие от задачи классификации. Привести примеры использования алгоритмов кластеризации.
 14. Парадигма Map Reduce. Описать принцип работы. Нарисовать схему. Перечислить слабые и сильные стороны. Обозначить области применимости. Привести примеры использования.
 15. Визуализация. Дать определение визуализации. Показать важность визуализации в аналитике больших данных. Привести примеры и инструменты для визуализации.
 16. Научные проблемы больших данных. Показать значимость проблем, актуальность, связь с областями математики и инженерии.
 17. OLAP и OLTP системы. Разница.
 18. Репликация и шардинг.
 19. Требования ACID. CAP-теорема, BASE архитектура
 20. NoSql. Классификация NoSql хранилищ. Их особенности. Примеры распределенных хранилищ.

7.2. Методические материалы, определяющие процедуру оценивания знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций.

Программой дисциплины в целях проверки прочности усвоения материала предусматривается проведение различных форм контроля:

1. «Входной» контроль определяет степень сформированности знаний, умений и навыков обучающегося, необходимым для освоения дисциплины и приобретенным в результате освоения предшествующих дисциплин.

2. Тематический контроль определяет степень усвоения обучающимися каждого раздела (темы в целом), их способности связать учебный материал с уже усвоенными знаниями, проследить развитие, усложнение явлений, понятий, основных идей.

3. Межсессионная аттестация–рейтинговый контроль знаний студентов, проводимый в середине семестра.

4. Рубежной формой контроля является зачет. Изучение дисциплины завершается зачетом, проводимым в виде письменного опроса с учетом текущего рейтинга.

Рейтинговая оценка знаний студентов проводится по следующим критериям:

Вид оцениваемой учебной работы студента	Баллы за единицу работы	Максимальное значение
Посещение всех лекции	макс. 5 баллов	5
Присутствие на всех практических занятиях	макс. 5 баллов	5
Оценивание работы на семинарских, практических, лабораторных занятиях	макс. 10 баллов	20
Самостоятельная работа	макс. 40 баллов	10
Промежуточный контроль по модулю	макс. 60	60
Итого		100

Неявка студента на промежуточный контроль в установленный срок без уважительной причины оценивается нулевым баллом. Повторная сдача в течение семестра не разрешается.

Дополнительные дни отчетности для студентов, пропустивших контрольную работу по уважительной причине, подтвержденной документально, устанавливаются преподавателем дополнительно.

Лабораторные работы, пропущенные без уважительной причины, должны быть отработаны до следующей контрольной точки, если сдаются позже, то оцениваются в 5 балл.

Знания, умения, навыки и опыт деятельности оцениваются по балльной системе на основе результатов тестирования, решения контрольных задач, участия в обсуждениях, представления рефератов. Оценки определяются с учётом индивидуальных особенностей студентов с максимально соблюдаемой объективностью вне зависимости от каких-либо внешних факторов (давления со стороны руководства, просьб и попыток подкупа).

Критерии оценивания:

"Зачет" - изложенный материал фактически верен, наличие глубоких исчерпывающих знаний в объеме пройденной программы дисциплины в соответствии с поставленными программой курса целями и задачами обучения; правильные, уверенные действия по применению полученных знаний на практике, грамотное и логически стройное изложение материала при ответе, усвоение основной и знакомство с дополнительной литературой;

- наличие твердых и достаточно полных знаний в объеме пройденной программы дисциплины в соответствии с целями обучения, правильные действия по применению знаний на практике, четкое изложение материала, допускаются отдельные логические и стилистические погрешности.

Практические задания выполнены на 60-100%.

"Не зачет"- баллов (оценка неудовлетворительно) - ответы не связаны с вопросами, наличие грубых ошибок в ответе, непонимание сущности излагаемого вопроса, неумение применять знания на практике, неуверенность и неточность ответов на дополнительные и наводящие вопросы». **Выполнено менее 60% практических заданий.**

Оценочные средства для инвалидов и лиц с ограниченными возможностями здоровья выбираются с учетом их индивидуальных психофизических особенностей.

- при необходимости инвалидам и лицам с ограниченными возможностями здоровья предоставляется дополнительное время для подготовки ответа на экзамене;
- при проведении процедуры оценивания результатов обучения инвалидов и лиц с ограниченными возможностями здоровья предусматривается использование технических средств, необходимых им в связи с их индивидуальными особенностями;
- при необходимости для обучающихся с ограниченными возможностями здоровья и инвалидов процедура оценивания результатов обучения по дисциплине может проводиться в несколько этапов.

Процедура оценивания результатов обучения инвалидов и лиц с ограниченными возможностями здоровья по дисциплине предусматривает предоставление информации в формах, адаптированных к ограничениям их здоровья и восприятия информации:

Для лиц с нарушениями зрения:

- в печатной форме увеличенным шрифтом,
- в форме электронного документа. Для лиц с

нарушениями слуха:

- в печатной форме,
- в форме электронного документа.

Для лиц с нарушениями опорно-двигательного аппарата:

- в печатной форме,
- в форме электронного документа.

Данный перечень может быть конкретизирован в зависимости от контингента обучающихся.

8. Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины

8.1 Основная литература:

1. Крутиков, В.Н. Анализ данных : учебное пособие / В.Н. Крутиков, В.В. Мешечкин ; Министерство образования и науки Российской Федерации, Федеральное государственное бюджетное образовательное учреждение высшего профессионального образования «Кемеровский государственный университет». - Кемерово : Кемеровский государственный университет, 2014. - 138 с. : ил. - Библиогр. в кн. - ISBN 978-5-8353-1770-7 ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=278426>
2. Жуковский, О.И. Информационные технологии и анализ данных : учебное пособие / О.И. Жуковский ; Министерство образования и науки Российской Федерации, Томский Государственный Университет Систем Управления и Радиоэлектроники (ТУСУР). - Томск : Эль Контент, 2014. - 130 с. : схем., ил. - Библиогр.: с. 126. - ISBN 978-5-4332-0158-3 ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=480500>
3. Базы данных в высокопроизводительных информационных системах : учебное пособие / авт.-сост. Е.И. Николаев ; Министерство образования и науки РФ, Федеральное государственное автономное образовательное учреждение высшего образования «Северо-Кавказский федеральный университет». - Ставрополь : СКФУ, 2016. - 163 с. : ил. - Библиогр.: с.161. ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=466799>

8.2.Дополнительная литература:

1. Туманов, В.Е. Проектирование хранилищ данных для систем бизнес-аналитики : учебное пособие / В.Е. Туманов. - Москва : Интернет-Университет Информационных Технологий, 2010. - 616 с. : ил., табл., схем. - (Основы информационных технологий). - ISBN 978-5-9963-

0353-3 ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=233492>

2. Добронев, Б.С. Численный вероятностный анализ неопределенных данных : монография / Б.С. Добронев, О.А. Попова ; Министерство образования и науки Российской Федерации, Сибирский федеральный университет. - Красноярск : Сибирский федеральный университет, 2014. - 168 с. : граф., ил. - Библиогр. в кн. - ISBN 978-5-7638-3093-4 ; То же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&>

9. Перечень ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины.

1. Lectures on scientific computing with Python. В свободном доступе: URL : <https://github.com/jrjohansson/scientific-python-lectures>
2. Python. The official Python web site. В свободном доступе: URL : <https://www.python.org/>
3. Программирование и научные вычисления на языке Python В свободном доступе: <http://ru.wikiversity.org/wiki/>
4. Пакет NumPy. Краткое введение: URL : В свободном доступе: <http://pyviy.blogspot.ru/2009/09/numpy.html>
5. Мерков, Александр Борисович. Распознавание образов: введение в методы статистического обучения / А.Б. Мерков; Рос. акад. наук, Ин-т систем. анализа.— Москва: УРСС=URSS, 2010.— 254с.:ил. (URL:<http://www.recognition.mccme.ru/pub/RecognitionLab.html/slbook.pdf>)
6. Доклад ЦРУ про большие задачи и большие данные: URL: <http://bit.ly/CRUbigdata>
7. Тренажер для освоения основ языка R: <http://tryr.codeschool.com/> от O'Reilly
8. Язык R: из учебной лаборатории — в мир больших данных. Леонид Черняк. osp.ru/os/2012/04/13015768/

10. Методические указания для обучающихся по освоению дисциплины

По курсу предусмотрено проведение лекционных занятий, на которых дается основной систематизированный материал, лабораторных работ, контрольной работы, зачета и экзамена.

Важнейшим этапом курса является самостоятельная работа по дисциплине с использованием указанных литературных источников и методических указаний автора курса.

Виды и формы СР, сроки выполнения, формы контроля приведены выше в данном документе.

Для лучшего освоения дисциплины при защите ЛР студент должен ответить на несколько вопросов из лекционной части курса.

В освоении дисциплины инвалидами и лицами с ограниченными возможностями здоровья большое значение имеет индивидуальная учебная работа (консультации) – дополнительное разъяснение учебного материала.

Индивидуальные консультации по предмету являются важным фактором, способствующим индивидуализации обучения и установлению воспитательного контакта между преподавателем и обучающимся инвалидом или лицом с ограниченными возможностями здоровья.

11. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине

- Проверка домашних заданий и консультирование посредством электронной почты.
- Использование электронных презентаций при проведении лекций и практических занятий.

Перечень необходимого программного обеспечения

1. Python,
2. R, R Studio.
3. Программы для демонстрации и создания презентаций («Microsoft Power Point»).

Перечень информационных справочных систем:

1. ЭБС Издательства «Лань» <http://e.lanbook.com> ,
2. ЭБС «Университетская библиотека онлайн» www.biblioclub.ru ,
3. ЭБС «Юрайт» <http://www.biblio-online.ru> ,
4. ЭБС «ZNANIUM.COM» www.znanium.com ,
5. ЭБС «BOOK.ru» <https://www.book.ru>.

12. Материально-техническая база, необходимая для осуществления образовательного процесса по дисциплине

№	Вид работ	Материально-техническое обеспечение дисциплины
1.	Лекционные занятия	Лекционная аудитория, оснащенная презентационной техникой (проектор, экран, компьютер/ноутбук) и соответствующим программным обеспечением (ПО) PowerPoint. ауд. 129, 131, А305.
2.	Лабораторные занятия	Лаборатория, укомплектованная специализированными техническими средствами обучения – компьютерный класс, с возможностью подключения к сети «Интернет», программой экранного увеличения и обеспеченный доступом в электронную информационно-образовательную среду университета. (лаб. 102-106.).
3.	Групповые (индивидуальные) консультации	Аудитория, (кабинет) – компьютерный класс
4.	Текущий контроль, промежуточная аттестация	Аудитория, приспособленная для письменного ответа при промежуточной аттестации.
5.	Самостоятельная работа	Кабинет для самостоятельной работы, оснащенный компьютерной техникой с возможностью подключения к сети «Интернет», программой экранного увеличения и обеспеченный доступом в электронную информационно-образовательную среду университета.